格致方法·社会科学研究方法译丛



刘军 译

Gary Goertz, James Mahoney

[美] 加里·格尔茨 詹姆斯·马奥尼 著

两种传承

社会科学中的定性与定量研究

A Tale of Two Cultures

Qualitative and Quantitative Research in the Social Sciences

格致出版社 上海人民出版社

格致方法·社会科学研究方法译丛

刘军 译

Gary Goertz, James Mahoney
［美］加里·格尔茨 詹姆斯·马奥尼 著

两种传承

社会科学中的定性与定量研究

A Tale of Two Cultures

Qualitative and Quantitative Research in the Social Sciences

格致出版社 上海人民出版社

图书在版编目(CIP)数据

两种传承:社会科学中的定性与定量研究/(美)格尔茨(Goertz, G.),(美)马奥尼(Mahoney, J.)著;刘军译.—上海:格致出版社:上海人民出版社,2016

(格致方法.社会科学研究方法译丛)

ISBN 978-7-5432-2621-0

Ⅰ.①两… Ⅱ.①格… ②马… ③刘… Ⅲ.①社会科学-定性分析 ②社会科学-定量分析 Ⅳ.①C3

中国版本图书馆 CIP 数据核字(2016)第 075264 号

责任编辑 王亚丽
装帧设计 路 静

格致方法·社会科学研究方法译丛

两种传承:社会科学中的定性与定量研究

[美]加里·格尔茨 詹姆斯·马奥尼 著
刘军 译

出 版 世纪出版股份有限公司 格致出版社
世纪出版集团 上海人民出版社
(200001 上海福建中路 193 号 www.ewen.co)

格致出版

编辑部热线 021-63914988
市场部热线 021-63914081
www.hibooks.cn

发 行 上海世纪出版股份有限公司发行中心

印 刷 苏州望电印刷有限公司
开 本 635×965 1/16
印 张 17.75
插 页 2
字 数 250,000
版 次 2016 年 6 月第 1 版
印 次 2016 年 6 月第 1 次印刷

ISBN 978-7-5432-2621-0/C·148 定价:45.00 元

我非常钦佩这部著作。在定性方法论领域，格尔茨和马奥尼是广受尊敬的创新者。作者精通各项论题，论点推群独步，论证清晰确凿，使本书受众广布。《两种传承》必将激起有益的争论；它对政治学方法论是一大贡献。

——戴维·瓦尔德内尔，弗吉尼亚大学

本书文笔优美、令人信服。这部杰作必将得到广泛阅读、学习和引用。

——科林·埃尔曼，锡拉丘兹大学麦克斯韦学院

这是社会研究方法论的一部重要著作！它清楚地论证了定量和定性研究的本质特征及其差异，为我们理解社会研究的不同取向和传统提供了新的视角，值得社会研究者认真阅读、深入思考。

——风笑天，南京大学

我赞同作者的结论：定量研究和定性研究存在差别，但地位同等重要。二者的互补而非互相排斥，增进了我们对社会现象的理解。

——张静，北京大学

序　言

本书将社会科学中的定量和定性研究作为独立的传承(cultures)加以分析。我们是在过去十年的教学和研究过程中得出这“两种传承”观点的。我们不断地发现,定性与定量研究者在方法论导向和研究实践方面都采取了不同的方式。我们也发现,在这两类研究者之间存在着误解和勉强的交流。当我们尝试厘清这些事实时,逐渐清楚地知道,定量研究和定性研究的传统展示出两种不同传承的全部特征,包括不同的规范、实践和工具箱。

笔者撰写本书的目的是:在学术上促进理解这些传承在哪些方面存在差异以及差异背后的原理何在。为此,我们将广泛涉猎几乎全体社会科学家都要面对的方法论主题,涉及一些重要的研究设计和数据分析问题。本书讨论的许多主题在很多研究方法类教材中都没有提及,也不会在任何一本关于定量或定性方法论的实用类书籍中找到。由于任何社科学者在设计和开展研究时都要思考一系列问题,因而,在阅读和使用本书时,可以视之为一部指南,它为这些问题提供了指导。

我们最初是在定性与多元方法研究所(Institute for Qualitative and Multi-Method Research)的教学过程中了解到了彼此的研究。近些年以来,很多学生入读该研究所,并为我们论证的两种传承提供了反馈,我们在此表示感谢。尤其要感谢研究所所长科林·埃尔曼(Colin Elman)为我们的年度研究计划腾出空间。美国政治学会“定性与多元方法专业委员会”(Organized Section on Qualitative and Multi-Method

Research of the American Political Science Association)为本书中几个章节的初稿提供了时事通讯和会议报道，特此致谢。

我们的第一版论证是一篇文章，发表在 2006 年《政治分析》(*Political Analysis*)上。感谢《政治分析》的编辑罗伯特·S.埃里克森(Robert S. Erikson)对第一版的推进。我们如果不发表该文，可能就不会受到鼓励去继续发现和探究定量和定性研究的差异。

我们在亚利桑那大学(University of Arizona)和西北大学(Northwestern University)为研究生开设了方法论课程，在教学过程中讨论了本书手稿的部分内容。我们正是通过与许多采取定量和定性方法的研究生的互动，才将很多主题列在重要的方法论议题列表上。另外，本书的很多内容都在美国、欧洲和拉丁美洲的一些工作坊及短期课程中有过报告。很多研究生参加了这些讲习班、工作坊和短期课程并提出了洞见，在此表示感谢。我们还要特别感谢凯伦尼莎·莫罕达莉(Khairunnisa Mohamedali)和克里斯托夫·纽伦(Christoph Nguyen)完成了对附录中所报告文章的调查。感谢西北大学、威斯康星大学和耶鲁大学的师生们对本书的报告发表的评论。

普林斯顿大学出版社的查克·迈尔斯(Chuck Myers)帮助我们获取审稿报告，使我们从中受益。查克的工作还推进了本书的出版过程。感谢格伦达·克鲁帕(Glenda Krupa)润色手稿。第 13 章的卷首语最初出现在约翰·福克斯(John Fox)出版的《应用回归分析和广义线性模型》(第二版)中(*Applied Regression Analysis and Generalized Linear Models*)(Los Angeles, CA: Sage Publications, 2008)，感谢世哲出版集团(Sage Publications Inc.)授权使用。

最后，下面这些同仁都提供了富有洞见的评论：迈克尔·鲍姆加特纳(Michael Baumgartner)，纳撒尼尔·贝克(Nathaniel Beck)，安德鲁·贝内特(Andrew Bennett)，珍妮特·博克斯-斯蒂芬斯迈耶(Janet Box-Steffensmeier)，贝尔·布劳米勒(Bear Braumoller)，戴维·科利尔(David Collier)，撒德·唐宁(Thard Dunning)，科林·埃尔曼(Colin Elman)，约翰·耶林(John Gerring)，杰克·利维(Jack Levy)，戴安

娜·卡皮斯泽韦斯基(Diana Kapiszewski),查尔斯·C.拉金(Charles C.Ragin),卡斯滕·施奈德(Carsten Schneider),贾森·西赖特(Jason Seawright),戴维·瓦尔德纳(David Waldner)和塞巴斯蒂安·扎娅(Sebastian Zaja)。我们深知同仁们不会都赞同本书中的每一个观点,但是我们希望,如果参与并讨论本书提出的观点,这本身就有助于推进社会科学中的定量和定性研究。

目　录

序言　/Ⅰ

1　引　论　/1
　1.1　为什么是两种传承？　/3
　1.2　两种传承的特征和比较　/6
　1.3　定性研究有何特色？　/10
　1.4　结论　/14

2　数学导引：向社会科学家选介逻辑学与集合论　/18
　2.1　导言　/18
　2.2　自然语言与逻辑学　/20
　2.3　带有二分类的必要与充分条件　/21
　2.4　模糊集的必要和充分条件　/28
　2.5　聚合　/32
　2.6　面对数据的模型　/34
　2.7　结论　/39

第一部分　因果模型与推断

3　先因后果与先果后因　/45
　3.1　导言　/45
　3.2　研究问题的类型　/47
　3.3　个案　/51
　3.4　结论：互补性与混合方法研究　/53

4 因果模型 /57
4.1 导言 /57
4.2 两种因果模型 /58
4.3 集合论模型和个案内分析：一种阐释 /61
4.4 因果关系的复杂性 /63
4.5 殊途同归 /65
4.6 结论 /68

5 不对称性 /73
5.1 导言 /73
5.2 对称与不对称模型 /74
5.3 集合论模型的非对称性解释实例 /75
5.4 结论 /82

6 休谟对原因的两种界定 /85
6.1 导言 /85
6.2 定量传统 /86
6.3 定性传统 /89
6.4 结论 /92

第二部分 个案内分析

7 个案内对跨个案的因果分析 /99
7.1 导言 /99
7.2 个案研究/小样本研究中的个案内分析对比实验/统计研究 /101
7.3 因果过程观测比对数据集合观测 /102
7.4 过程追踪检验对比统计检验 /105
7.5 结论 /109

8 因果机制与过程追踪 /115
8.1 导言 /115
8.2 机制和因果推断 /116
8.3 多元方法中的过程追踪对比定性研究 /122

8.4 结论 /125

9 反事实 /134
9.1 导言 /134
9.2 因果推断的基本问题:不同的回答 /135
9.3 建构反事实 /137
9.4 结论 /142

第三部分 概念与测量

10 概念:定义、指标与误差 /149
10.1 导言 /149
10.2 对特征及指标的界定 /150
10.3 误差 /154
10.4 结论 /158

11 意义和测量 /163
11.1 导言 /163
11.2 语义学变换和集合-成员函数 /166
11.3 次要变异原理 /168
11.4 隶属函数和量表类型 /170
11.5 结论 /172

12 语义学、统计学和数据转换 /176
12.1 导言 /176
12.2 标准化比对意义保留 /177
12.3 对数转换—模糊集转换 /179
12.4 数据转换的基本原理 /181
12.5 结论 /184

13 概念的对立面与类型学 /188
13.1 导言 /188
13.2 针对概念对立面的对称及非对称的研究 /189
13.3 重叠对比排他的类型学 /192

13.4 语义学和名义类型学 /194
13.5 结论 /198

第四部分 研究设计与普遍化

14 个案选择与假设检验 /205
14.1 导言 /205
14.2 基于因变量进行选择 /206
14.3 选择个案的策略 /210
14.4 结论 /216

15 概括 /223
15.1 导言 /223
15.2 定性的概括 /224
15.3 集合论概括和 2×2 表 /227
15.4 统计模型、完美预测项和集合论概括 /229
15.5 控制变量和完美预测项 /231
15.6 结论 /232

16 范围 /239
16.1 导言 /239
16.2 针对因果异质性问题的模型内回应 /241
16.3 为什么利用范围条件？ /244
16.4 范围和经验检验 /249
16.5 结论 /252

17 结论 /257
17.1 差异概要 /258
17.2 社会科学中的方法论多元主义 /262

附录 /264
术语对照表 /268
译后记 /270

1 引　论

本书将探究社会科学领域中定性研究和定量研究传统之间的关系，特别关注政治科学和社会学。为此，我们将识别这些传统在哪些方面有别。这些对比出现在方法论的诸多领域，从研究问题的类型，到数据分析的模式，再到推断方法中皆有对比。我们认为，这些差异是系统、连贯且彼此相关的，因此，论及不同的定量和定性研究范式是有意义的。

我们把定量和定性研究传统看做可供选择的传承。每一种研究传统都有自己的价值、信仰和规范。每一种传统都关系到特定的研究程序和实践。每一种传承内部的交流通常是顺畅和富有成效的。然而，不同传承之间往往难以交流，误解频仍。当来自某个研究传统的学者向另一个传统的成员发表洞见时，这种建议通常被视为是无用和不当的。这两种研究传统之间并不协调，存在着误解、怀疑和互挫，有时这些标志着定性和定量研究者之间的交锋。究其核心，我们认为，社会科学中的定性定量之争确实是不同研究传统的碰撞。

与所有的传承一样，定量传承和定性传承也不是铁板一块（关于"传承"概念的精彩论述，参见 Sewell，2005）。它们是被松散地整合起来的传统，包含着内部的矛盾和争论。随着时间的推移，构成这些传承的特殊目标和实践已在演变，持续至今。现在，这两种传承不再彼此隔绝，相反，它们开始相互渗透和跨界。虽然如此，这两种传承还是相对连贯的意义和实践系统。它们表现出许多容易识别的价值、信念、规范和程序等特征。

通过强调定量和定性研究的差异，本书因而有别于金、基欧汉和维巴的著作《社会科学中的研究设计》（*Designing Social Inquiry：Scientific Inference in Qualitative Research*）。他们的著名论断是"定量

与定性研究传统仅在风格上有差异，并且这些差异在方法论和实质意义上都不重要”(King, Keohane and Verba, 1994:4)。他们相信，这两种研究传统共享单一推理逻辑，该逻辑推理主要可以通过统计分析的规范加以总结。他们识别出来的两种传统之间的差异涉及一些表面特质，特别是使用数字还是文字。

我们拒绝如下假设，即基于统计规范的单一推理逻辑可以同时指导定量和定性研究。我们也不相信定量定性是围绕着数字对文字的使用而分立的。相反，我们考虑的是研究的基本取向上的差异，例如学者是否主要运用个案内分析(within-case analysis)来推断个体个案(如定性研究者所做)，或是否采用跨个案分析去推断总体(如定量研究者所为)。我们甚至建议，最好诉诸两个备择性的数学基础来理解这两个传统：定量研究建立在推断统计学(即概率论和统计学理论)基础上，而定性研究则(经常暗含地)根植于逻辑学和集合论。依据这些对比性的数学基础来看待这两个传统，有助于理解本书讨论的许多差异。

在指出基本分歧的过程中，我们的目标并不是将定量和定性研究范式分裂开来。相反，本书致力于促进不同范式的研究者之间的交流和合作。我们认为，相互理解必然建立在对差异的认知与欣赏的基础之上，包括理解各自对照性的优缺点。当问题要求研究者所追求的目标兼具定性与定量范式的特性时，我们提倡跨界和混合方法的研究。与此同时，我们也尊重在每一种范式中进行的研究，从不认为这种研究具有内在的劣势。在社会科学中，定量研究、定性研究和混合方法研究都占有一席之地。

从本书引出的一个教训是，询问定性研究或定量研究是否优于对方，这不是一个有益的问法。金、基欧汉和维巴(King, Keohane and Verba, 1994:5—6)也声称：“定量研究和定性研究都不优于对方。”然而，他们之所以得到此结论，仅仅是因为他们认为当统计分析不可行时，迫不得已才必须使用定性方法。[①]相比之下，我们认为，定量和定性

---

① 如他们所说：“某些主题可以利用定量数据进行统计上的假设检验，进而得到有意义的表述，而社会科学家感兴趣的许多主题都不能如此表述，所以我们不希望鼓励完全使用定量技术。”(King, Keohane and Verba, 1994:6)

方法适于不同的研究任务，其设计也是为了达成不同的研究目标。至于选用定量还是定性方法，这与恰好可用的数据无关。当然，对于某些研究目标而言，定量方法比定性方法更适用，而对于其他研究问题来说，定性方法比定量方法更恰当。不过，因任务不同还可能出现这样的情形，即分析者必须两种方法都使用才能达其目标。对于许多复杂的研究项目来说，由于完成其目标要求研究者同时使用两种传统的取向和特有优势，因此，将定量和定性技术相结合的混合方法研究是关键所在。

与研究其他文化的某些人类学家一样，我们也力图理解一些研究实践，并对它们保持某种中立。我们的目标主要是描述性的，基本上不是规范性或规定性的。当然，这两种传统中的方法都不能免遭批判。然而，我们认为，对方法进行批评和重构最好在一个给定传统的内部进行。因此，统计方法论学者是最有资格提升统计方法的研究者，而定性方法论者占据着提升定性方法的最佳位置。我们发现，现存的"跨行"(cross-cultural)批判(比如定性研究者批判定量研究)并不恰切，因为他们忽视了那个传统中的基本目标和研究目的。从一类传统的角度看可能存在问题，从另外一类传统的视角看却可能完全可以理解。

我们在讲述这两种传承的故事时，结局总要考察鲜为人知、暗含的定性假定和实践如何有别于众所周知、精心编码的定量研究。这种考察路数是如下事实的副产品，即与定性方法相比，定量方法在社会科学中有更明晰、系统的发展。定量方法更为人所知，在大多数社会科学领域，定量传承毫无疑问在两种传承中更占优势。正因如此，我们泼墨去探讨定性方法。然而，通篇采用的分析路数仍然是要澄清两种传统的特色，同时也避免招致不满的比较。

1.1 为什么是两种传承？

金、基欧汉和维巴认为，有一个单一的推理逻辑(即一种基本传

承）可以刻画所有社会科学的特征，既包括定量研究，也包括定性研究。而另一种“多重传承”(many cultures)观点则认为，定量和定性传统是异质性的群组，每个群组内部都存在许多变体和子传承(subcultures)。的确，与任何一种传承类似，每个范式又都可以进行或大或小的划分。例如，从历史上讲，在统计学范式内有一个大的分割出现在经典的频次学派和统计分析的贝叶斯研究之间(例如，参见 Freedman, 2010; Jackman, 2009)。对于那些可能赞同例如频次-贝叶斯之争的学者来说，他们中间又可能有其他小分工，他们关注的问题有诸如混合效应模型的效用或者在一个统计模型中应该包含的自变量个数。

同样，定性范式也有许多分工。可能最大的区分是学者之间的差异，有的学者坚持广义的行为研究传统，关注的核心是因果推理，有的学者则采用各种解释性方法。这两大派又有各自的子类。例如，推崇实现因果推断目标的定性学者可能不赞成某些特定的工具——如反事实分析或者定性比较分析(Qualitative Comparative Analysis, QCA)的相对重要性。同样，在解释性研究的阵营中，信奉克利福德·格尔茨(Clifford Goertz, 1973)式诠释性分析的学者和提倡批判理论及后结构主义方法的学者之间也存在差异。

我们给出的两种传承分析路数与金、基欧汉和维巴的一种传承分析路数有某些相似性，尤其是因为我们的关注点集中指向因果推断和普遍概括性。我们所探讨的方法和技术都意在用来得出有效的科学推断。使用科学方法得出有效的科学推断，这是高于一切的工作，它将本书讨论的两种研究传统结合起来了。

我们对因果推断的关注有一个后果，即定性范式内部的一些重要流派将不在分析之列。特别是，在我们的两种传统论证中并没有描绘诠释性研究。这些研究路数通常不关注因果分析；它们更专注于其他研究目标，如阐释行为的意义或者批判权力的运用。诠释性传统有自己的主导规范和实践，基本上不同于本书所研究的定量和定性范式。人们当然可以另撰专著，专门考察诠释性传承与我们所讨论的“因果推断”传承在哪些方面可做对比。这样一部书会公开社会科学的各个部

分内部存在的认识论和本体论的根本冲突。然而在本书中，我们关注那些在认识论和本体论的一些根本议题上达成一致——包括认为因果分析在理解社会世界方面居于核心——的学者。①

至于为什么专注于这两种研究传统是有意义的，则有各种原因。一方面，定性定量之分已经植入几乎每一位社会科学者的词典中，并且它被作为一个共同的参照点来区分不同类型的研究工作。几乎所有的研究者都会谈及定量对定性研究，尽管他们可能不是按照同样的方式来理解这种对比。甚至像我们这样的学者都感觉到，“定量”和“定性”这样的标签根本不足以捕捉到两种传统之间的最显著差异，但是我们仍然被迫使用这个术语。

此外，社会科学家已经将自己（正式和非正式地）归入定量和定性研究的阵营之中。在政治科学中有两类方法论专业委员会，一类是代表定量方法论的政治学方法论专业委员会（Section on Political Methodology），较新的一类是定性与多元方法研究专业委员会（Section on Qualitative and Multi-Method Research）。在社会学中，方法论专业委员会（Section on Methodology）主要代表定量方法，而我们所讨论的各种定性方法则与比较和历史社会学专业委员会（Section on Comparative and Historical Sociology）相关。一些领先的培训机构同样也反映了两种传统的分野：政治与社会研究校际联盟（Interuniversity Consortium for Political and Social Research，ICPSR）提供的几乎都是定量训练，而定性与多元方法研究所（Institute for Qualitative and Multi-Method Research，IQMR）则关注定性和混合方法研究。

本书的目的不是把定量研究者转变成定性研究者，或者反之。然而，我们的确努力让更多研究者能够理解这两种研究传承的规范和实践及其基本原理。我们认为，在社会科学中，克服定量定性之分显然是

① 我们决定本书不考虑诠释性研究，这并不证明我们认为此类研究在社会科学中没有立足之地。事实上，广义地讲，我们关于两种传承的论证实乃关于描述和解释的操练。我们试图阐明两种相对一致的研究传承的实践及相关的意义。因此，诠释性分析者虽然在我们描述的定性传承所代表的研究传统中找不到他们的研究传统，但是会发现他们的传统中的许多工具会用在我们的分析中。

有意义的，它让我们更好地理解两种传统之间的方法论差异，以及差异的原因何在。

1.2 两种传承的特征和比较

在讨论定量和定性传统之时，我们要利用各类数据资源，并关注某些类型的实践而非其他类型。在本节中，我们简单描述用什么方法来刻画和比较这两种传承。

1.2.1 资料类型

我们对研究实践的刻画来自三种资料。第一，我们依赖于有关定量和定性方法论的文献。在一个既定传统中要清晰地显现所使用的研究技术及其背后的基本原理，方法论者在此方面经常做出漂亮的成果。就定量范式而言，我们大量使用了在统计学、计量经济学和定量社会科学领域中著名学者的教材。我们的展示大量援引了有关 Neyman-Rubin-Holland 模型及相关的“潜在结果架构”(potential outcomes framework)方面的文献(如 Angrist and Pischke, 2009; Berk, 2004; Freedman, 2010; Morgan and Winship, 2007)。本书也参考了社会科学中关于实验研究的文献。乔瓦尼·萨托里(Giovanni Sartori)、亚历山大·乔治(Alexander George)和戴维·科利尔(David Collier)等学者对定性范式做过一些研究，我们的讨论建立在这些研究中的“经典准则”(classic cannon)基础之上。另外，我们还借鉴了查尔斯·拉金(Charles Ragin)的许多洞见。在每一章的最后，我们都推荐一些书和文章，感兴趣者可以深入阅读这些文献，进一步探讨该章讨论的差异。

第二，我们利用一些范例性的定量和定性研究来展示每一章中讨论的差异。这些研究不仅可用作个案，还可以作为来源提供有关两种传承中的特色性实践的洞见。其中的一些范例关注的主题对于两种传

统来说都很重要，例如民主研究。从每一种传承的角度来考虑在范例性研究中所处理的同一个主题，可使我们更加生动地展示驱动这两种传承的不同类型的问题和方法。然而，我们同时有一个重要的论点，即某些主题在一种传承中比在另外一种传承中更容易得到探究。因此，我们的一些例子并没有跨越两种传承。

第三，我们也从政治科学和社会学主流期刊的大量研究性论文中进行抽样和编码。附录中汇总了所编码的条目及结果。这个大样本意在代表好的研究，正如出现在主流的政治科学和社会学期刊中的文章所定义的那样。样本也为有关主导性的研究实践的推广提供了进一步基础。例如，当我们做出声明，宣布定量研究者在他们的统计模型中经常包含几个控制变量的时候，该声明就建立在我们调查结果的基础之上。

1.2.2 显在和潜在的实践

我们的讨论关注定量和定性范式中的一些占主导的方法论实践。通常情况下，在探讨定量研究的时候，我们关注那些遵循来自方法论文献中建议的显在的实践。定量研究方法和程序经常是明确设定的，并且定量研究者往往明确地遵循着这些精心表述的方法论思想。

然而，我们将在很多方面讨论定量传统中通常隐含的假定和程序。比较定量研究与定性研究就是为了引人关注两个传统中潜在的规范和实践，否则它们可能不被注意。例如，通过考虑许多定性方法的不对称性假定(assymetry assumption)，就可以比较清晰地看到大多数定量方法在多大程度上潜在地假定了对称关系。系统地比较这些范式，有助于揭示那些经常被认为是理所当然的研究实践。

在对待定性研究方面，我们更加关注一系列潜在的程序和技术。通常情况下，与定量方法相比，定性方法的运用是相当不显见的。实际上，在这个阶段，研究方法的潜在使用可能被看做定性研究在传承上的特征。为了描述这个研究传统，我们必须重构定性研究者在做研究时所使用的程序。我们的重构有赖于广泛阅读定性研究成果，包括努力

对诸多定性研究文章进行系统地编码。此外，我们描述的实践与其他方法论教材明确地处理和编码的定性研究实践是一致的（例如，Brady and Collier，2010；George and Bennett，2005；Ragin，1987）。然而，由于定性方法的使用不具有系统性，该传统的某些特征会不可避免地引起争议。我们描述的定性研究中的主流实践在某些领域中可能有争议，本书尝试指出这些领域。

1.2.3 典型的实践、最佳的实践与可能的实践

对于任何传统来说，典型的实践和所谓的最佳实践之间都有张力（正如被一流方法论专家所识别的那样）。在社会科学中，对“最佳实践”的认定通常有很大争议。在一个既定的传统内部，方法论专家就赞成还是反对特定的研究程序进行着辩论。这些辩论指出了在定性和定量方法论领域内部存在着不同的子领域。例如，在定量方法论领域，大多数研究都尝试用观测数据做因果推断，提倡实验的研究者对此却持有重大保留意见。

人们在方法论上争论着什么能够构成最佳实践，本书不参与这些争论。取而代之，我们的兴趣点在于描述研究者的实际所为，我们会关注定量和定性传统中的典型研究实践，将它们界定为坊间出版的重要成果。我们所考察的实践是进行社会科学分析的标准工具。它们被广泛（尽管不普遍）地认为是可以接受的，并适合于进行描述性及因果性推断。的确，从更广的学术共同体角度看，这些典型实践都是“好的实践”，因为采用这些实践的研究是有影响力的（在积极意义上），并且通常出现在同行评议的顶尖期刊和最权威的出版社所出版的著作中。因此，学者常常会生产出被整个学术共同体评为是最好的研究，我们的分析就关注这样的一些实践。

在讨论两种传承实践之间的差异时，我们不否认定量研究者可能会模仿定性实践或反之。然而，我们在这里关注的是实际的实践，而不是所谓的“可能的实践”。例如，可能去改变统计研究的 Neyman-Rubin-Holland 模型的构型来处理定性研究中显著的问题，诸如对充

要条件的分析。然而,在定量传承中,对充要条件的研究并不是一件必做之事,事实上,在实践中也从来没有做过。同样,与定性范式相关的集合论分析数学模式可用于分析一个总体中的平均因果效应。但是,据我们所知,没有任何一位社会科学研究者为了达到此目的而使用过这些方法。我们的观点无非是,某些系列的工具会自然带来某些类型的实践而不是其他实践。当一个人在冥想采用哪些方式将一种传承中的工具扩展到另一种传承中易于完成的事项的时候,这些扩展不是自然的,并且通常是纯粹假设性的。

定量范式会更直接地刻画在受到高度关注的研究中通常使用的一些做法,因为定量方法在一些有名的教材中都有相当明确的展示。应用研究者从这些教材中学习方法,并且通常尽可能地严格遵照其中的规则。当然,教材之间并不总是一致的,它们随时间在变。尽管如此,它们为定量传统中的许多共享规范和实践提供了依据。

在定性传统中,这种情形是比较随意的。尽管谈论起菜谱统计学(cookbook statistics)比较容易,我们却从未听闻有人用过"菜谱定性分析"(cookbook qualitative analysis)这种表述。尽管存在着许多定性方法教材(著作),但是学生们不要指望在定性方法课程中学到单独的一套核心技术。究其原因,部分在于在定性研究内部存在着两类学者之间的分立,一类学者集中关注因果推断,另一类则使用诠释性方法论。同样,正是由于定性研究对方法的运用是潜在的,才使得与定量范式相比,定性研究远远不是标准化的。

不管怎样,如果我们关注定性研究的因果推断学派,就可以区分出并讨论一套潜在的且相当常用的做法。这些做法在许多定性研究名家的著述中都可以看到,在有关定性研究的影响力大的方法论著作中也有描述,如布雷迪和科利尔(Brady and Collier, 2010),乔治和贝内特(George and Bennett, 2005),格林(Gerring, 2007)以及拉金(Ragin, 1987)。

我们希望,通过考察出现在高水平期刊和著作中的这些典型实践,学者可以就如何做更好的研究提出新观点。其实现方式可以多元。一种可能是,来自一个既定传统的学者可能发现,源自其他传统的某些

思想能有助于在他们自己的传统中开展实践。例如，对于定量研究者来说，如何提升测量的效度，有关概念形成的定性分析或许可以提供新的洞见。相反，当定性研究者在做出自己的描述性推断的时候，或许可以受益于关于测量误差方面的大量统计学文献。这些都显示出跨行学习(cross-cultural learning)的可行性，这个主题会在本书中多次提及。

另外一种可能性是，一种给定的实践在研究者自己的传统内部很常见，但由于这种实践与研究者自己认为是最好的实践并不一致，因而使研究者惊讶。例如提倡 Neyman-Rubin-Holland 模型的定量方法论者可能惊讶地认识到，实际上该模型能够影响社会科学研究的程度是有限的。在定性研究方面，提倡中等样本量的定性比较分析(medium-N QCA)研究的学者可能发现，对于大多数定性研究中的因果推断来说，个案内分析仍然是核心的基础，这一认识会令他感兴趣。我们认为，要想赞同、批判和提升占主导地位的研究实践，就必须对这些实践有很好的理解。本书为提升这种理解提供了基础。

1.3 定性研究有何特色?

由于定性方法经常被潜在地运用，因此，我们希望标识出两种重要的工具，并相信这两种工具可以刻画区别于定量研究的定性传统的特征。第一种工具是个案内分析(within-case analysis)技术，如过程追踪法，这些技术在有关政治科学中的定性方法的许多主流研究中被强调过，最著名者可能包括亚历山大·L.乔治和安德鲁·贝内特(Alexander L.George and Andrew Bennett)所著的《社会科学中的个案分析和理论发展》，亨利·E.布雷迪和戴维·科利尔(Henry E.Brady and David Collier)主编的《反思社会研究:多元工具与共享标准》。第二种工具是逻辑学与集合论，它们几乎散布在所有主要的定性技术(包括个案内分析)中，并且经常与查尔斯·拉金(Charles Ragin, 2000, 2008)的研究关联在一起。

1.3.1 个案内分析

区分定量研究和定性研究的一个常用的方法是关注样本量。这很自然涉及统计分析的大样本(Large-N)研究和定性探究的小样本(small-N)研究。金、基欧汉和维巴(King, Keohane and Verba, 1994)在讨论定性研究时,用很大篇幅来分析定性研究的小样本问题,或分析当缺乏足够样本时,应用常规的统计学方法进行推断的难点何在。这种分析路数遵循的是一脉相承的研究,即根据自由度问题来考察定性方法论(Lijphart, 1971;Campbell, 1975)。

然而,某些样本量相当大的研究也被认为是定性研究,其他样本量相当少的研究却使用主流的统计学方法(例如,可参见 Collier, Brady and Seawright, 2010:178—179)。这个事实表明,尽管小样本与定性研究相关联,但是它不能界定这样的研究。在界定定性研究时,最重要的是使用个案内分析,它要求对特定的个案有广泛的了解,这种使用因而有利于解释为什么大多数定性研究有一个小样本。定性研究者可以选择一个小样本,因为他们的核心推断方法——个案内分析——要求有一种个案导向的分析,而这很难用一个大样本来完成。

如果人们关注个案内分析,视之为定性研究的核心特性,那么这种把定性研究和小样本问题联系起来的思想就不应在讨论之列。显然,定性研究体现出自己的因果分析方式。个案内分析涉及要使用一些特定片段的数据和信息来对个案进行推断。这些个案内的观测可能成为"确凿证据"(smoking gun),它明确地支持或反对一个给定的理论。在这种情况下,从自由度问题角度来考虑定性方法论是没有用的。

与定性研究相比,统计学方法显然几乎都是跨个案分析的工具。我们可以在实验方法中看到这一点,实验法往往被作为定量范式中因果推断的黄金标准。一个实验涉及两组对照性的被试者,一组得到处理,一组被控制。因果推断根本上是围绕这种跨个案的比较建立起来的。人们不会试图去解释,比如那些受到处理的特殊个案会发生什么。实验法的设计不是为了告诉我们对于任何一位被试者来说,这种处理

是否带来结果。尽管观测式分析在很多重要的方面(比如,研究设计)不同于实验法,但是它们共享一种基本的跨个案的因果推断法。

1.3.2 逻辑学和集合论

当定性研究者在口头上明确表达他们的理论时,他们很自然地在使用逻辑学的语言。我们将这种现象称为定性研究和逻辑学之间关系的"茹尔丹先生"(Monsieur Jourdain)* 性质。①定性研究者使用逻辑学语言,但是他们往往没有意识到这个事实。然而,为了系统地描述定性研究的实践,有必要使这种逻辑的潜在使用变得明确化和形式化。

有关充要条件的思想居于定性研究实践的核心。在定性传统中,这些类型的条件被隐含地用于形成众多研究假设。它们是概念形成的定性方法、个案选择的定性探究以及几乎所有假设检验的定性方法的核心成分。假设检验的定性方法是围绕充要条件建立起来的,包括穆勒的求同法和求异法、重要的过程追踪检验[如箍筛检验(hoop tests)和确凿证据检验(smoking gun tests)]以及所有的定性比较分析(QCA)模式。我们的看法是,如果不考虑必要性和充分性的观念,定性研究和方法论就不能得到充分的编码和理解。

有一长串术语直接或间接地表明,研究者在提出假设时使用了逻辑学资源。为了表达这样的因果观念,即"X 对于 Y 来说是必要的",研究者要利用一些术语和表述,如"只要……就"、"对……来说是本质的、不可少的、必需的、必要的","阻碍、否决、阻止","是……的必要条件"以及"能够、允许、许可"等。在利用逻辑学来表述因果关系的性质方面,其中的一些表达是很明确的和直接的,如"只要 X……就有 Y"。其他的表达就不那么明确,尽管还算清楚,如"X 对于 Y 来说必不可

---

* 茹尔丹先生是法国剧作家莫里哀(Molière)的剧作《贵人迷》(*Le Bourgeois Gentilhomme*)中的主角。该剧作写的是一位巴黎富商茹尔丹,他一心想当贵族,被别人玩弄却还自以为乐,当不成本国贵族就当一个土耳其的假贵族。剧中一个人物自豪地声称:"一个既有钱又长得体面的正人君子比一个肮脏而穷酸的贵族可强得多。"——译者注

① 在莫里哀的笔下,当茹尔丹得知自己的诗歌老师用散文来言说时非常激动。

少”或者“并不是 X 阻止了 Y”。

类似地，有很多术语可以表示研究者理解了 X 是 Y 的充分条件。在这种情况下，研究者用诸如“保证、确保”，“总是跟随……”，“不可避免地导致”以及“带来、生成、产生”等词和表达。其中的一些术语还是会比其他表述（“X 产生 Y”）更直接地表明一种充分关系（例如：“Y 总是跟随着 X”）。

一旦人们对于自然的逻辑语言的使用变得敏感，就会在社会科学文献中随时见到它们。对于定性研究者（或者就此而言，对于任何研究者）来说，利用上述列举的一种或多种术语来表达一个口头的理论是非常平常的事情。我们遇到过数百个有关必要或充分条件的假设的例子。①在所研究的学术问题中，这些假设不是偶然被提出的；相反，它们居于正在提出的陈述的核心（有关必要条件假设的 150 个例子，参见 Goertz，2003）。

对逻辑学和集合论的应用超出了对假设的表述。为了运用与乔瓦尼·萨托里（Giovanni Sartori，1970）有关的经典定性方法来定义一个概念，需要建构这样一系列条件，即单独看它们是概念的必要条件，整合起来就成为概念的充分条件。萨托里传统中的定性研究者“自然”会接受逻辑学，将它作为一种架构来思考概念化问题。②同样，当人们使用穆勒的求同法来“消除”一个假设时，就潜在地假定该假设有一个必要条件。正如我们将看到的那样，甚至主要的过程追踪检验——如“箍筛检验”和“确凿证据检验”——都可以基于关于充分性和必要性的观念来预测。

① 这个名单包含许多著名的比较社会学家，如斯考切波（Skocpol，1979：154），穆尔（Moore，1996：418），吕金迈耶、斯蒂文和斯蒂文斯（Rueschemeyer，Stephen and Stephens，1992：270）以及政治科学的一些最著名的比较政治学家，如奥唐奈和施米特（O'Donnell and Schmitter，1986：65），林茨和斯特潘（Linz and Stepan，1996：61），利瓦伊（Levi，1988：144）。在国际关系中，几乎所有的学界领军人物都会（暗含地）提出这些类型的假设，包括（新）实在主义者如华尔兹（Waltz，1979：121，参见 Levi 和 Thompson（2010）给出的扩展的讨论），自由制度主义者如基欧汉（Keohane，1980：137），杨和奥谢恩科（Young and Osherenko，1993）以及诸如温特（Wendt，1992：396）和芬尼莫尔（Finnermore，1996：158）这样的社会建构主义者。

② 当然，萨托里（Sartori）非常清楚地知道自己的研究的逻辑基础。

本书将通篇讨论定性研究的程序和方法借鉴哪些方式来使用逻辑概念。实际上,数理逻辑及其同类的集合论在社会科学中还不为大众所知,因此,本书需要在序言中简要介绍它们。现在,我们希望强调的是,逻辑学和充分性与必要性的观念不仅仅是查尔斯·拉金开发的QCA技术中所使用的工具。确切地说,在过去几十年间,它们还是定性研究者在他们的多元方法研究中一直在暗中使用的资源。

1.4 结论

在本书的最后,我们希望读者对“定量”和“定性”这两个术语有所质疑。我们将讨论这两个范式之间大量重要的差异,但是这些差异不能用这些术语很好地识别出来,特别是如果这些术语被理解为意味着数字对文字的话,就更不能这样来识别了。

在结论部分,我们总结了本书给出的许多对比。我们在两种传承之间提供了一个共含25个差异的清单。尽管某些差异——如个案内分析对跨个案分析以及统计学对逻辑学——是我们论证的核心,但我们并不认为任何单个对比能推动所有其他对比。相反,我们的结论是,每一种传承都是由许多不同的规范和实践组成的,它们相对一致地共同作用在一起。

展望未来,读者可以有不同的方式阅读本书。尽管我们已经试图将各章组成几个连贯的部分,读者也未必按照任何固定的顺序去阅读各章。每一章都尝试完整独立成篇。因此,读者可以依据兴趣挑选专题去阅读,跳过一些章节也不会有障碍。在利用了充要条件观念的一些方法中,需要运用逻辑学和集合论知识。对于在这些方法方面缺乏基础的读者来说,数学导引部分有选择地介绍了这些知识。有了导引的铺垫,我们应该考虑这两种传承如何以相当不同的方式来看待和解释相同的数据,尽管这些方式都是同样正当的。

参考文献与推荐读物

Angrist, J.D., and J.-S.Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Berk, R. 2004. *Regression Analysis: A Constructive Critique*. Newbury Park: Sage Publications.

Brady, H.E., and D.Collier, eds. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition. Lanham, MD: Rowman Littlefield.

Campbell, D. 1975. "Degrees of Freedom and the Case Study." *Comparative Political Studies* 8:17—93.

Collier, D., H.E.Brady, and J.Seawright. 2010. "Sources of Leverage in Causal Inference: toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E.Brady and D.Collier. Lanham, MD: Rowman Littlefield.

Finnemore, M. 1996. "Constructing Norms of Humanitarian Intervention." In *The Culture of National Security: Norms, Identity, and World Politics*, edited by P.Katzenstein. New York: Columbia University Press.

Freedman, D. 2010. *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*. Cambridge: Cambridge University Press.

Geertz, C. 1973. *On the Interpretation of Cultures*. New York: Basic Books.

George, A.L., and A.Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Gerring, J. 2007. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Goertz, G. 2003. "The Substantive Importance of Necessary

Condition Hypotheses." In *Necessary Conditions: Theory, Methodology, and Applications*, edited by G. Goertz and H. Starr. New York: Rowman & Littlefield.

Jackman, S. 2009. *Bayesian Analysis for the Social Sciences*. New York: Wiley.

Keohane, R. 1980. "The Theory of Hegemonic Stability and Changes in International Regimes, 1967—1977." In *Changes in the International System*, edited by O. Holsti, R. Siverson, and A. George. Boulder: Westview.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Levi, M. 1988. *Of Rule and Revenue*. Berkeley: University of California Press.

Levy, J., and W. Thompson. 2010. "Balancing at Sea: do States Ally against the Leading Global Power?" *International Security* 35: 7—43.

Lijphart, A. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65:682—693.

Linz, J., and A. Stepan. 1996. *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-communist Europe*. Baltimore: Johns Hopkins University Press.

Moore, B. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Morgan, S., and C. Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

O'Donnell, G., and P. Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions About Uncertain Democracies*.

Baltimore: Johns Hopkins University Press.

Ragin, C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Rueschemeyer, D., E.Stephens, and J.Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.

Sartori, G. 1970. "Concept Misformation in Comparative Politics." *American Political Science Review* 64:1033—1053.

Sewell, Jr., W.H.2005. "The Concept(s) of Culture." In *Logics of History: Social Theory and Social Transformation*, by W. H. Sewell, Jr. Chicago: University of Chicago Press.

2 数学导引：向社会科学家选介逻辑学与集合论

我的深深抱怨是，政治科学家明显缺乏逻辑学，甚至基础逻辑学训练（当然也有例外）。

——乔瓦尼·萨托里（Giovanni Sartori）

2.1 导言

在本导引中，我们将讨论逻辑学与集合论的一些核心观念，这些观念可以引导我们在接下来的章节讨论定性研究。①我们无意对逻辑学与集合论领域进行某种全面介绍，因为这需要用一本书的篇幅。相反，这里的讨论是有选择的，即我们关注与定性方法论相关的观念，尤其是涉及必要与充分条件的观念。

本章的数学导引之所以理当有之，原因在于，定性研究内在的主要数学取向——逻辑学与集合论——并不为大多数社会科学家（包括大多数隐含地使用这些理论的定性研究者）所熟知。尽管其他学术领域（如哲学、系统工程、数学、人工智能和计算机科学）中出现了许多关于逻辑学与集合论的著作，但实际上却没有任何面向社会科

① 本书考察的"逻辑学"与"集合论"基本上指的是同样的数学领域。有时候，本书自然会讨论和运用逻辑学符号，有时候运用集合论资源则更为容易。有关逻辑学与集合论的不同的隐喻式基础，参见拉考夫和努涅斯（Lakoff and Núñez，2000）。

学的相关著作。[①]简而言之，本书需要有一章数学导引，因为定性研究的数学取向在社会科学中几乎从未被讨论过。

该讨论也在定性与定量研究之间给出首个主要对比：它们建立在不同的数学传统基础之上。定量研究侧重于与统计学和概率论相关的数学工具。大多数社会科学家对这些工具都很熟悉，因为它们在研究方法课程上被广泛地讲授，也明确地用于定量研究中。有人甚至说，大多数社会科学家都假定统计学-概率论是社会科学的数学。然而我们认为，定性研究通常或明或暗地以集合论与逻辑学为基础，并且，如果人们想要比较定性研究与定量研究，就必须理解这些数学工具本身。[②]

对于我们的讨论来讲，特别重要的是在模糊集合分析中使用的数理逻辑。尽管哲学家需要学习逻辑学——正像社会科学家需要学习统计学一样，但模糊集合分析却不是关于逻辑学的哲学教材中的主题。相反，它经常出现于一些应用领域，如工程学、计算机科学与专家系统。专家系统的设计者使用模糊集合数学来制造"智能"机，其范围涉及从洗衣机到电梯再到摄像机（McNeill and Freiberger, 1994）。尽管可以证明模糊集合数学在现实社会中有实际功用，但作为数据分析的一种形式工具，它直到最近才被引入社会科学中（如 Smithson, 1988; Ragin, 2000, 2008）。

本导引面对的是定量与定性研究者。对于定量研究者而言，本章描述了一种非统计性的研究传承的数学基础。正如逻辑学课程提供的工具不会出现在任何统计学课程中一样，本章所讨论的步骤也不同于统计学的主要方法。对定性研究者而言，本导引揭示了他们通常隐含地使用的数学取向。

---

① 查尔斯·拉金（Charles Ragin）和其他学者（如 Schneider and Wagemann, forthcoming）的著述在定性比较分析（Qualitative Comparative Analysis, QCA）传统上明确地讨论了逻辑学和集合论。然而，仍然没有被广泛使用的教材可为社会科学家提供导引。

② 在一些技术类期刊，如《模糊集和系统》（*Fuzzy Sets and Systems*）中，关于统计学与集合论的对比是一个流行的主题。

2.2 自然语言与逻辑学

当定性学者系统地口述其理论时，他们会非常自然地运用逻辑学语言。在上一章，我们涉及定性研究与逻辑学之间关系的"茹尔丹先生"性质。定性研究者谈及的是逻辑学语言，但他们通常没有完全意识到这一事实。①

运用逻辑学和集合论来表述的理论通常具备两个成分。第一，这种理论将概念视为集合或范畴，其中的个案（或观测项）可以具有成员资格，也可能包括部分程度的成员资格。在日常语言中，诸如民主、发展与战争这样的概念被视作范畴，而其中的特定个案可能具备或不具备其成员资格，或者具备一定程度的成员资格（参见 Lakoff，1987）。如下所述，在概念分析方面，逻辑学与集合论保留了许多日常语言。

第二，在两个或多个概念之间有某些有假设性的、用逻辑学术语来构想的关联，其中要用到有关必要性和/或充分性，或与之对等的超集/子集关系的思想。尽管某些研究者可能一打眼就认为必要性和/或充分性观念对社会科学研究来说是不适当的，本书却认为情况并非如此。正如导言部分所说，相当多的学者在提出假设、建构概念、选择个案和检验假设时，都使用了必要性/充分性思想。

我们可以比较定性传承中的自然逻辑学语言和定量传承中的概率论与统计学语言。几乎所有人都熟悉后者，因为几乎每个人都在统计学课程上学习过，并从运用统计方法的期刊文章上了解到它。例如，为人熟知的公式包括：

- 在 X 出现或发生的情况下，Y 发生的概率增加（或降低）。
- 在 X 出现或发生的情况下，Y 平均而言发生的概率增加（或降低）。

① 形式建模者也会论及必要与充分条件。在一些数学定理（如在经济学或博弈论领域）中，提供必要和/或充分条件是很常见的（例子参见 Goertz，2003b）。然而，当实施经验检验时，则由统计学范式接手过来了。

X 与 Y 之间关系的函数形式有多种，这取决于所使用的特定统计模型。例如，函数形式可以是线性的常规最小二乘法(OLS)回归、曲线形状的概率单位(probit)或 logit 以及其他统计模型的对数线性形式，如贸易引力模型(gravity model of trade)。当使用概率模型时，函数式可能仍然未被设定：$P(Y \mid X) \neq P(Y \mid \neg X)$。

在两种传承中，学者有时没有注意到统计假设与逻辑假设的任何区别。当他们陈述自己的假设时，甚至可能在这两种假设之间反复切换，好像两者可以互换一样。例如，当华尔兹主要使用自然语言的逻辑来陈述假设时，他在某一点指出："团体越小……一些成员——即越多的成员——将越可能为团体的利益而行动……一个单位的相对规模越大，它越有可能将自己的利益与系统的利益认同起来。"(Waltz, 1979：198)在此，我们看到了典型的定量传承语言。然而，他在最后一章中回归到了逻辑学，并主张"在当前条件下，广泛的国际合作只有在美国的领导下才是可能的"(1979:210)。

在整个经验社会科学领域中有一个基本问题，即如何将口头的理论形式化以便诉诸检验。人们能否真实地将用逻辑学术语表述的口头理论转化为与常规的统计方法一致的形式？同样，人们能否将依据概率论与统计学术语表述的某种理论重塑为一种建立在必要与充分条件基础上的理论？这些问题提出了我们所谓的转译问题(translation problem)。为了探究跨范式的转换为什么会成为一个问题，我们需要更细致地考察逻辑学与集合论。

2.3 带有二分类的必要与充分条件

2.3.1 集合论与维恩图

有时候，定性分析者明确地使用有关必要性与充分性的逻辑学术语，但同样是这些研究者，他们却很少明确地运用集合论术语。不过，

命题逻辑与集合论是密切相关的，通常可以互换使用。这种转译工作如下：

- “X 是 Y 的一个必要条件”等价于“Y 是 X 的一个子集”。
- “X 是 Y 的一个充分条件”等价于“X 是 Y 的一个子集”。

图 2.1 运用维恩图（Venn diagrams）来展示这种思想。为了向这些图中加入内容，可以使用几个简单的范畴实例。在图 2.1a 中，令 Y 表示未通过基础逻辑学课程期末考试的学生集合，X 表示缺席基础逻辑学期末考试的学生集合。那么，缺席期末考试的集合中的成员就是未通过课程考试的学生集合的一个充分条件。X 是 Y 的充分但不必要条件，因为还有其他课程不合格的原因（例如，课程论文不及格）。这种通向同一结果的多种途径的观念被称作“殊途同归”（equifinality）。

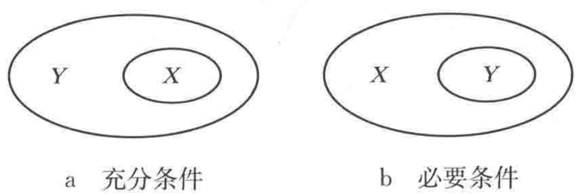


图 2.1　用维恩图展示的必要条件与充分条件

为了向图 2.1b 增加简单的内容，可以假设 Y 表示所有怀孕个体的集合，X 表示所有女性个体的集合。那么，某成员属于女性集合，是其属于怀孕个体集合的必要条件。因此，X 不是 Y 的充分条件，因为如维恩图所明示，还有许多女性未怀孕。

集合论语言可以通过多种表达形式隐含地出现在数据分析中，其中的一种形式是“所有 X 都是 Y”。如果学者说“所有 X 都是 Y”，那么她也是在说 X 是 Y 的一个子集。相反地，如果她说“所有 Y 都是 X”，那么 Y 就是 X 的一个子集。从集合论的视角看，在“所有 X 是 Y”与“所有 Y 是 X”之间有很大差别。但是，怎样将这种思想转译成定量语言呢？这两种陈述都意味着 X 与 Y 之间的某种关系；对于二者来说，

X 与 Y 之间通常存在某种密切的关系。但是，如何将集合论的陈述转译为概率论或统计学的语言，这仍然不是一目了然的。原则上，我们大概可以进行这样的转换（例如，某些转译可参见 Goertz, 2003a：第 10 章；Cioffi-Revilla; 1998; Seawright, 2011），但这并非自然而然或轻而易举能够完成的。

定性研究者在尝试将一种线性相关或统计关联转换为必要与充分条件的时候会面临难题，上述转换问题类似于此难题。虽然有许多方法可以进行这种转换（例如 Eliason and Stryker, 2009），但这些方法是有局限的和勉强的。正如我们的两种传承隐喻所示，在一种传承中是显见和容易之事通常在另一种传承中是有问题和困难的（尽管不是不可能的）。

2.3.2　2×2 表

学者在描述必要或充分条件的时候，最常见的方法可能是运用 2×2 表。表 2.1a 和表 2.1b 展示了这些条件在二分类别中是如何出现的。为了有助于记住必要条件与充分条件的差异，我们可以使用前文展示的例子：在表 2.1a 中，成为女性（$X=1$）是怀孕（$Y=1$）的必要条件；在表 2.1b 中，缺席终考（$X=1$）是未通过考试（$Y=1$）的充分条件。

表 2.1　运用 2×2 表来展示必要条件与充分条件

a　必要条件

		X = 0	X = 1
Y	1	0	N_2
	0	N_3	N_4

b　充分条件

		X = 0	X = 1
Y	1	N_1	N_2
	0	N_3	0

当运用 2×2 表的术语表达时，说 X 是 Y 的必要条件则意味着三种相互关联的情形：(1)“没有 $Y=1$ 会是 $X=0$”；(2)“所有 $X=0$ 都是 $Y=0$”；(3)“所有 $Y=1$ 都是 $X=1$”。因此，用我们的例子来说，这就意味着：(1)没有怀孕之人不是女性(即为男性)；(2)所有非女性(即男性)都不怀孕；(3)所有怀孕之人都是女性。[①]2×2 表的关键特征是，单元格($\neg X$, Y)是空的($\neg X$ 读作“非 X”)。事实上，可以将该单元称为一个 2×2 表的必要条件单元(necessary condition cell)。与之类似，在表 2.1b 中，充分条件单元是(X, $\neg Y$)，因为要想充分条件成立，该单元格必须为 0。

表 2.2　一个充分条件的例子：民主式和平

	非民主二方组	民主二方组
和平	1 045	169
战争	36	0

资料来源：Russett，1995：174。

为了将这种讨论建立在社会科学基础之上，让我们考察表 2.2 中给出的民主式和平(democratic peace)涉及的数据。当评价民主式和平理论的时候，涉及的个案都是由国家组成的二方组(dyads)(即一对国家)。所关注的主要结果是和平，它被看成是一个二分类范畴，和平的对立面是战争(该观点将在“概念的对立面与类型学”一章中进行论辩，但这里用此概念是很适合的)。“民主二方组”(democratic dyads)作为原因的要素，也是一种二分范畴。只有当两个国家都是民主政体时，二者才构成一个民主二方组。只要其中一个或两个国家不是民主政体，该二方组就是“非民主二方组”。

从逻辑学的角度讲，2×2 表中的数据是充分条件的一个绝佳个案。具体而言，“民主二方组 X”是“和平 Y”的充分条件。这一点可以在该表中清晰地看到，其中民主二方组与战争对应的单元是空的，所有

① 一些研究者会声称，“非女性”不等于男性。正如我们将在第 13 章所讨论的，一个范畴的否不等于该范畴的对立面。

的民主二方组个案都处在和平单元格中。

当然，人们可以针对如表 2.2 这样的 2×2 表计算统计量。结果将因使用的统计量不同而有所差别。例如，针对表 2.2 的数据计算卡方统计量 χ^2，可得到卡方值 5.80，显著性水平为 0.02。斯皮尔曼和皮尔森相关系数值 τ_b 为−0.07，在 0.006 水平上是统计显著的。①发生比率(odds ratio)则更有启示意义：它在 0.08 的水平上是非常显著的（发生比率如果等于 1，意味着没有关系；其值越接近 0 或远远大于 1，则表明是非常显著的结果）。如此看来，用 logit 分析计算得到的发生比率表明，民主二方组实质上比非民主二方组更讲和平。然而，这些标准的统计量都不能挖掘到这样的事实，即该数据与充分性关系是完全一致的。

关于必要条件的陈述总是能够被转换为关于充分条件的陈述（反之亦可）。在从必要性转换到充分性（或者反之）时，只要简单地否定所分析的范畴，就可以实现这种转换。例如，如果 X 是 Y 的充分条件，那么 $\neg X$ 是 $\neg Y$ 的必要条件。在民主与和平的个案中，可以将一个重要发现表述如下：民主二方组的缺失是战争发生的一个必要条件。

在逻辑上，人们通常不会混淆必要条件与充分条件；对于统计方法来说，同样会出现这样的关系。事实上，不考虑数据的分布是否呈现为必要条件或充分条件，计算 2×2 表的诸多相关系数结果会完全相同。确实如此，因为标准的相关系数计算方法假定有对称关系，其设计并不是用来探测和概括非对称关系的。

当然，对于许多 2×2 数据集的个案来说，用对称的相关测度来分析其分布是适用的。对于双变量相关分析来说，这的确为真。在这样的分析中，个案都集中在两个对角线单元上（例如，如果是正相关，个案则集中在左下与右上的单元中）。当一位头脑中装备了逻辑学与集合论的学者面对这样一个对称的数据集时，他或她可能将其视为具有某些必要性和某些充分性。这样来解释数据没有任何错误。关键在于，当使用逻辑学与集合论时，根据对称相关来思考这些关系是欠妥当的

① 严格地讲，当存在零单元格时，许多 2×2 表的相关系数（如 χ^2）是不能计算的。然而，大多数统计软件包具有标准的补丁程序，允许计算这些统计量。

(尽管不是不可能)。

关于 2×2 表的这种讨论展示了所谓的罗夏测验原理(Rorschach Principle)。进行罗夏测验时要向人们展示模糊不清的图像,并要求他们作出解释。对社会科学家而言,数据可以发挥同样的作用。毕竟,数据不足以支持理论,这是科学的一项核心原则。人们可以考察相同的数据,却合法地看到不同的东西。没有哪一种看待数据的方法是绝对正确的(尽管并非所有的看待数据方法都同样有用)。2×2 表说明,人们在数据中如何看到不同的模式,这取决于各自的研究目的。

2.3.3 真值表

在以逻辑学为基础的研究中会采用真值表,它类似于在统计分析中所分析的数据集,尽管它们在令人感兴趣的某些方面有所不同。表 2.3 给出了一个真值表。正如在定量的矩形数据集合中一样,列表示变量。由于这是一个"经验"真值表,它也包含第四列来表示经验观测值,这些值对应于真值的每一种配置(configuration)(详见 Ragin, 1987:第 7 章)。

表 2.3 经验真值表

X_1	X_2	Y	N
1	1	1	5
1	1	0	0
1	0	1	3
1	0	0	10
0	1	1	0
0	1	0	7
0	0	1	0
0	0	0	12

尽管大部分列看上去与统计分析中的矩形数据集相类似,但是行却大不相同。在一个统计数据集中,行是观测值。相比之下,真值表中

的行却是真值的配置。这里列举了逻辑上所有可能的配置，使得变量数决定着行数。行数与观测值的数目无关。某一行，比如第一行，是这样一个逻辑陈述：$X_1=1$ 且 $X_2=1$ 且 $Y=1$。数据与该陈述可能一致或不一致（表 2.3 的数据是一致的，无矛盾）。

作为一种方法论，定性比较分析（Qualitative Comparative Analysis，QCA）可以描述和分析嵌入在诸如本真值表中的逻辑关系（Ragin，1987，2000）。它运用布尔代数数学（如布尔极小化和蕴涵[Boolean minimization and implication]）将逻辑表述还原为更简化的形式。我们希望强调的是，变量值的配置（并非单个变量）构成了分析的核心。

通常情况下，一种配置就是对两个或多个变量值的组合（或交集），这些变量组合在一起，成为某个结果的充分条件。构成一个配置的逐个个体变量值是通过逻辑"与"（AND）关联在一起的。这些个体变量值通常是"充分非必要条件的必要非充分部分"（INUS conditions）（或者它们可能是必要条件；Mackie，1965，1980）。[1]INUS 条件是这样的一些变量值，即它们的每一个值对于所研究的结果来说都既非必要也非充分。相反，它们是诸多变量值的一种总配置的本质（即非冗余的）要素，这些变量值构成了结果的充分条件。因此，当变量值的某种组合足以产生一个结果时，单个变量值就或者是必要条件，或者是 INUS 条件。如果是后者，即殊途同归总是存在的，这意味着通向同一结果之途不止一条。

我们可以利用统计方法来分析表 2.3 中的数据。但是，不能按照给定的方式来分析数据。例如，需要将行转换成观测值。由于有 37 个观测值，统计数据集将有 37 行。进行这一步时，在真值表中那些没有个案的逻辑配置（即第二、五、七行）事实上将不做进一步考察。统计分析者可能探讨每一个自变量（X_1 和 X_2）与 Y 之间的协变性。相比之下，采用定性比较分析的学者可能将数据中令人感兴趣的模式总结如下：

① 首字母缩写 INUS 源自麦凯（Mackie，1965：246）的如下所述："对于一个结果来说，所谓的原因是该结果得以成立的一个非必要但充分的（unnecessary but sufficient）条件中的那个非充分但必要的（insufficient but necessary）部分。"

- $X_1=1$ 是 $Y=1$ 的一个潜在的必要条件。
- $X_1=1$ 且 $X_2=1$ 是 $Y=1$ 的一个潜在的充分条件组合。

考察 X_1 与 Y 之间以及 X_2 与 Y 之间的统计共变特征，这根本就不对。与利用基于逻辑学的方法来汇总结果相比，这种分析既非更正确也非更错误。它无非揭示了数据的一个不同方面罢了。比较理想化的是，给定了研究的问题与目标后，为了提请注意数据的最相关特征，人们会有办法并倾向于从多个视角来考察数据。

2.4 模糊集的必要和充分条件

亚里士多德的逻辑学和我们熟知的清晰集合论(crisp-set theory)将概念或变量看成是二分的，即个案或者是其成员，或者不是。某类别中的成员属性通常表示为数量 1(即 $X=1$)，成员属性的缺失被表示为数值 0(即 $X=0$)。模糊逻辑的一个特色在于，允许个案在一定程度上隶属于某类。完全的成员资格仍然标记为数值 1，完全的非成员资格值是 0。但是，个案可以有 0 到 1 之间的任何值(例如，$X=0.75$，$X=0.33$，$X=0.10$)，这取决于它们在多大程度上属于所感兴趣的类别。如此看来，一个给定的个案属于"战争"类别的成员资格值可以是 0.5，这表明该个案居于该类内、外的程度是一样的。

我们可以在 2×2 表中讨论必要条件和充分条件，与之类似，也可以针对 X 和 Y 之间的连续散点图讨论必要和充分条件。就模糊集分析来说这是可能和自然的，因为个案在集合中的成员资格可以用 0 到 1 之间的值来连续地测量。当针对两个类别来绘制模糊集数值的图像时，必须使 X 轴和 Y 轴的范围在 0 到 1 之间连续变化。

虽然大多数学者没有经过任何特殊的训练就用充分条件和必要条件来解释 2×2 表，但对于模糊集散点图来说情况并非如此。在 2×2 情况下，我们要注意到与充分条件和必要条件有联系的空单元。那么

对于连续编码的模糊集来说，我们要注意什么呢？

图 2.2a 和图 2.2b 分别表示的是，当将个案绘制到一个模糊集散点图上的时候，必要条件和充分条件看起来是什么样子的（参见 Ragin, 2000）。它们表现为“三角形”散点图。右下方的三角形散点图是一个必要条件，左上方的三角形散点图是充分条件。我们可以把散点图看做对我们所分析的 2×2 表的扩展。在这些表中，3/4 的单元格有数据。有数据的这三个单元格构成了一个类似三角形的形状。这样看来，如果描绘这些表，就可以得到如图 2.2 中的散点图。

更正式地说，针对一个必要条件（X）的模糊逻辑规则是，所有个案在 X 上的模糊集值必须大于或等于其在 Y 上的值：

X 是 Y 的必要条件：对于所有的 i 来说，$x_i \geqslant y_i$，x_i，$y_i \in [0, 1]$。

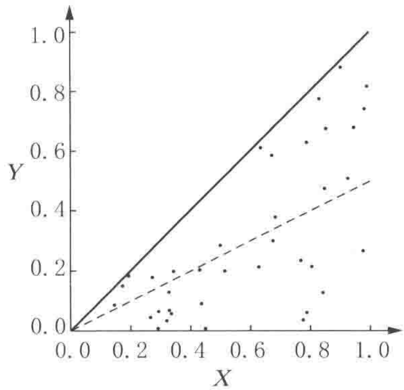
其背后的观念是，X 要想成为 Y 的一个必要条件，就要求一个个案隶属于 X 的程度必须不小于其隶属于 Y 的程度，否则，说 X 是 Y 的必要条件就没有什么意义。举个例子，如果一个个案隶属于 X 的程度很小（例如 0.1），却完全是 Y 的成员（即 1.0），那么就不能说该个案隶属于 X 的程度对于它属于 Y 的程度而言是必要的，否则就会出错（除非做出其他特殊的假定）。

相比之下，当论及模糊逻辑中的充分条件时，其意思是说所有个案在 X 上的模糊集合值（fuzzy-set values）将小于或等于它们在 Y 上的值：

X 是 Y 的充分条件：对于所有的 i 来说，$x_i \leqslant y_i$，x_i，$y_i \in [0, 1]$。

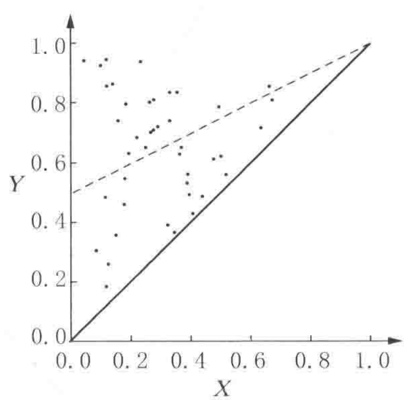
这里的观念是，如果 X 是 Y 的充分条件，那么一个个案成为 Y 中成员的程度必须不小于其成为 X 中成员的程度。例如，如果一个个案完全是 Y 的成员（即 1.0），但是仅在较小程度上成为 X 中的成员（如 0.1），

那么说它是 X 的成员就潜在地足以说明它可以成为 Y 的成员，这不会有问题。



注：虚线是 OLS 线。

a　必要条件



注：虚线是 OSL 线。

b　充分条件

图 2.2　模糊逻辑：必要条件和充分条件图示

如果回到前文讨论的维恩图（即图 2.1），我们会在明确集（即二分的集合）和模糊集之间的关系上看到如下一致性：

必要条件：$Y \subseteq X \equiv y_i \leqslant x_i$

充分条件：$Y \supseteq X \equiv y_i \geqslant x_i$

甚至符号看起来都相似，这并不是偶然的。

从统计学视角出发，对散点图实施罗夏测验就是有益的。对于刚刚修完一两门回归课程的人来说，这些散点图有哪些特征会引人注意呢？

1. 适度拟合。人们可以画一条线穿过数据，发现在 X 和 Y 之间有一个清晰但适度的关系。

2. 斜率。图 2.2a 中的 OLS 斜率与图 2.2b 中的 OLS 斜率是一样的。

3. 异方差性。围绕 OLS 线的方差明显不是常量。

让我们从两种传承的角度依次考察这三点。首先，从模糊逻辑的角度看，图 2.2 中的散点图可以看做完美的拟合(perfect fits)。对于一个必要条件来说，当所有的观测值都位于或者低于 45°的对角线(根据上述的必要条件定义)的时候，就发生了完美拟合。同样地，对于一个充分条件而言，在所有的观测值都位于或者高于 45°对角线的时候，就发生了完美拟合。相比之下，在 OLS 回归模型中，当全部点都恰好落在 OLS 线上的时候，才发生完美拟合。

其次，从统计学角度讲，图 2.2 的两个散点图中 X 和 Y 的关系基本相同。截距固然不同，但那些参数很少是人们的兴趣点。相比之下，在模糊集分析中，人们就不能得出结论认为这两个散点图代表着相同的关系：在一张图中发现 X 对 Y 是必要的，在另一张图中却发现 X 对 Y 是充分的，可以认为这两个发现完全不同。这一点与我们在 2×2 表中看到的完全一致。定性传承将一个表看做必要条件，另一个表看做充分条件，定量传承却将这两个表解释成代表同样的关系。

当然，从集合论角度看，这些都是不对称的散点图，因此，它们对于展示集合论的潜在功用是非常有用的。然而，如果人们开始分析的是这样的两个对称散点图，即其中的点都随机分散在直线周围，那么定量

视角可以强调指出其中的细微差别和不同点，而这些是用集合论工具难以表达的。再次强调的是，我们的观点并不是说一个视角正确而另一个视角错误。确切地说，我们认为，这些视角可以注意到数据的不同特征。集合论模型对于分析不对称关系来说是非常有用的（虽然它们也可以用来研究对称关系）。

再次，在面对异方差性数据的时候，标准的统计学反应是对变量进行转换，使之达至常量的方差。在定量传统中，这种转换对于有效的统计推断来说可能非常重要。相比之下，定性研究者通常并不进行这种转换。对他们来说，转换一个变量通常意味着改变其意义。除非显示这样的转换能够增加被分析的基本概念的意义，否则变换就是不恰当的（参见第 12 章）。

2.5 聚合

方法论中有一个基本问题是选择什么方式来聚合（aggregation）数据。为了便于说明，让我们将聚合定义为一个函数 f，人们用这个函数将各个 X 组合在一起得到 Y：$Y=f(X_1, X_2, \cdots)$。函数 f 可以采取很多不同的形式，这取决于人们采用什么方法和假定。逻辑学和统计学都有各自关于聚合程序的独特的、默认的假定。

在统计学中，一种常见的聚合技术就是加权和。如下所示的一般线性模型就是一个好例子：

$$Y=w_0+w_1X_1+w_2X_2+\cdots \tag{2.1}$$

这里的 Y 是各个 X 的加权和。当权数都是 $1/N$ 时，就得到加权聚合的一个特例，即均值。

另一个常见的统计聚合形式是交互项：$X_1 * X_2$。通常情况下，交互项应该是诸如方程（2.1）这样的更大的加权线性聚合项的一部分。然而，如下所示的包含一个纯交互项的聚合模型也不是不可能有的：

$$Y = X_1^{w_1} * X_2^{w_2} * X_3^{w_3} * \cdots \quad (2.2)$$

在实践中，这种方程会被转化为如方程(2.1)那样的对数线性模型。在实施聚合的时候，定量传承当然不限于加法和乘法。但是，出于某些良好的实用性和统计学原因，大多数统计模型都是加权总和或对数线性的加权和。

对于或明或暗地使用逻辑学和集合论的学者来说，他们使用的标准聚合技术有哪些？我们对表 2.3 中真值表的讨论给出了大量的例子，每一行有一例，它们都有同样的基本形式。例如，第一行可读作：

$$X_1 = 1 \text{ 与 } X_2 = 1 \text{ 与 } Y = 1 \quad (2.3)$$

逻辑“与”(AND)将诸多条件联系在一起，变成诸多条件的一个总组合。在布尔代数中，逻辑“与”经常写作乘法符号(*)。

如果人们希望进一步比较逻辑学和统计学，就可以尝试在逻辑“与”和统计乘法之间转译。这两个数学运算可以写成同样的形式：$Y = X_1 * X_2$。如果 X_i 是二分变量，这个类比就是有效的。当超出二分变量时，关于“与”和乘法的类比就开始无效。例如，如果将变量理解成在[0, 1]上连续编码的模糊集，就可以针对方程(2.3)追问聚合程序。在逻辑学中，用逻辑“与”(AND)(即布尔代数中的乘法符号)计算一个个案的成员资格属性的标准规则是利用 X 的最小值：

$$Y = \min(X_1, X_2, X_3, \cdots), \ X_i \in [0, 1] \quad (2.4)$$

因此，有了逻辑“与”，一个个案隶属于 Y 的成员资格值就等于其在集合 X 中的最小值。例如，如果在集合 X 中，最小的模糊集值是 0.1，那么该个案隶属于 Y 的程度值就是 0.1。对于二分变量来说，运用逻辑“与”的最小值也很管用。

更明确地说：

乘法不同于运用最小值。

诸如此类的差异居于两种传承之争的核心。统计分析中的乘法和集合

论分析中的逻辑“与”类似，但是它们的数学过程却不同。此外，一个过程并不是内在地优于另一个过程。它们只是不同的数据聚合方法罢了。在某些理论的或实质的情景下，学者可能有理由优先选择其中的一种聚合模式而不选择另外一种。但是，并不存在先验的理由相信某种聚合模型应该被首选。

两种传承之隐喻也允许人们可以运用对方传承中的聚合技术。原则上讲，没有什么道理可以阻止统计学传承使用模型：$Y=\beta_0+\beta_1X_1+\min(X_2, X_3)$。同样，加权总和与乘法在模糊逻辑系统内也是可能的。然而，这两种情况至少在社会科学中是不自然的，也不常用。

上文讨论了乘法和逻辑“与”的关系，那么对于加法来说，也有类似的关系吗？逻辑“或”(OR)扮演了在利用逻辑式进行聚合时的加法角色。在布尔代数中，逻辑“或”被写作加号(+)，这又一次强调了类比性。在利用逻辑“或”来计算个案的成员资格程度的时候，一般规则是运用个案在集合 X 中的成员资格的最大值：

$$Y=\max(X_1, X_2, X_3, \cdots),\ X_i\in[0, 1] \tag{2.5}$$

例如，如果集合 X 中的最大模糊集合值是 0.85，那么个案成为集合 Y 的成员的程度值就是 0.85。这个方程也适用于二分变量，因此，我们可以将它看成是针对逻辑“或”的一般聚合程序。

需要再次明确的是，加法与运用最大值是不同的。虽然它们可能产生同样的结果，但通常不会如此。再强调一次，我们的论断不是说某个数学过程优于另一个。已知它们在各自的研究传承中是标准聚合程序的情况下，我们仅试图关注它们的差异。

2.6 面对数据的模型

如何评价经验数据和理论、模型、假设之间的“拟合”，这是一个基本的科学活动。在社会科学内部，这类评价通常由统计分析来进行。

例如，就总体的模型拟合而言，R^2 是一个经典测度（尽管在当代定量社会科学中总体拟合测度不再被认为是非常重要的）。要想考察单个变量的拟合度，人们需要追问数据是否"支持"关于该变量的假设。例如，人们会考察单个变量的因果效应和统计显著性。

但是，在逻辑学或集合论中如何评价一个模型或原因要素？什么样的标准可以评价模型的"拟合度"或某个既定的原因要素的"重要性"？本节将探究这些问题。不管是否用到定性比较分析（QCA），我们讨论的观念都与定性研究相关。

2.6.1 拟合与一致性

在评价一个理论或模型的拟合度时，一项基本要求是该理论或模型能够对数据做出明确的预测（预测本身可以采取不同的形式）。如果预测是明晰的，人们就可以问：数据与模型在多大程度上一致？在统计模型中，高度一致的数据会靠近或落在模型所描述的直线上（假定它是一个参数模型）。学生在第一门统计学课堂上，通常学习如何用肉眼看双变量散点图，考察是否存在一条拟合数据的线。

在集合论分析中，"一致性"这个术语被学者用作一种拟合测度，用它来分析关于必要或充分条件的假设（Ragin，2008）。不论用二分测度还是模糊集编码来测量必要和充分条件假设，都可以明确地预测数据的模式。就 2×2 表这个简单例子（参考表 2.1）来说，一个必要条件假定会预测在（$\neg X$，Y）单元格中应该没有个案。人们可以想象各种场景使得该单元不包含个案，而这只占一小部分。这些个案是所研究假设的"反例"。在图 2.1 中，如果有理由认为 N_2、N_3、N_4 的值大，那么在（$\neg X$，Y）单元格中的这"小部分"个案便意味着一致性不是百分之百的，但是仍然足够大到可以支持假设。数据可能显示，$Y=1$ 行中 95％的个案（这里选统计学的一个流行标准）位于（X，Y）单元格中（回想一下，必要条件的一种定义就是所有的 $Y=1$ 都是 $X=1$）。基于此，人们可以声称数据与必要条件假定是相当一致的。

这种基本观点同样适用于模糊集假设。如图 2.2 所示，如果数据

与充分条件假设是完美一致的，它们必然落在或高于45°对角线上。当观测值移动到低于45°对角线时，一致性会减弱。这里的窍门是设计一个公式（参见 Ragin，2008），用它来总结这些观测值不一致的程度。这类似于在 OLS 回归中，用离差平方和除以总变异：它总结了各个观测值同 OLS 线相比的不一致程度。

简而言之，有一些经过了良好设计的方法可用来评价基于逻辑的假设和模型在多大程度上与数据拟合。虽然在此过程中所用的具体公式不同于统计学中的共识，但是其道理是相同的，即将模型的预测值与实际数据进行对比。如果它们接近，模型或理论通常就得到了数据的支持。

2.6.2 评价重要性：覆盖度和琐碎性

有些假设是用必要条件来表述的。就这样的假设而言，人们有时候提出这样一个关切的问题，即某个因素可能是一个结果的必要原因，但它却是琐碎的原因（trivial cause）。[①]关于这个关切的问题，乔治·唐斯（George Downs）提供了很好的陈述："对于任何现象来说，都存在着无数个必要条件。例如，不可否认，所有的军队都需要水和重力才能运转，但是这些普遍项的贡献不大。"（Downs，1989：234）

在对琐碎性的这种批判内部包含着这样一种观念，即必要（或充分）条件因其"重要性"而变。在统计模型中有几种方式可以决定某个给定变量的重要性。例如，直线的斜率是某个变量的重要性的一个指标：斜率越陡，通常意味着变量越重要。为了测量充分条件和必要条件的相对重要性，也可以制定严格的标准。实际上，唐斯的论证提供了一个有用的出发点。他暗指某些必要条件无足轻重，因为它们一直存在。这种观测表明，必要条件的相对频次可能与它们的重要性有关。

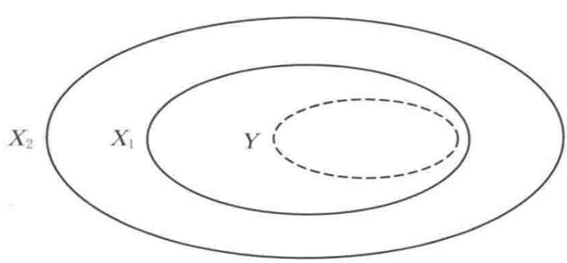
我们可以用集合论工具来系统地思考这个问题（Ragin，2008；Goerts，2006）。X 如果是 Y 的必要条件，则 X 是 Y 的一个父集（su-

① 有意思的是，虽然这些关切对于必要条件来说经常出现，但是学者并不经常提出它们，或者至少不会针对充分条件明确地提出它们。

perset)(换言之，即 Y 是 X 的子集)。这个定义并没有告诉我们 X 集合与 Y 集合之间关系的相对规模。作为一项规则，如果必要条件 X 在与 Y 集合的关联中成为后者的一个较小父集，则 X 会更重要；也就是说，X 集合会随着与 Y 集合趋于有完美的重叠而变得愈发重要。

另外一个思考该问题的方法是，询问一个"琐碎的"或不重要的必要条件的对立面是什么。显见的答案是，一个必要条件同时也是结果的充分条件。用集合论的术语来说，一个必要条件在多大程度上接近于充分条件，这可以表达为子集 Y 在多大程度上"包含"或填充父集 X。在极限情况下，当集合 X 和集合 Y 完全相等时，X 便是 Y 的充分必要条件，其重要性也达到最大。

请看一例，现在考虑这样一项发现，即地心引力和威权政体(authoritarian regime)的存在是社会革命的必要条件(即如果不存在地心引力或威权政体，一个社会就不会经历社会革命)。在这两个必要条件中，威权政体这个因素显然更加重要。然而，如何用集合论来表达这个事实呢？正如图 2.3 所示，我们可以做这样的展示，即拥有地心引力的个案集合包含所有社会的全部人群，因此，就必要条件来说，集合 Y"涵盖"集合 X 的程度达到最小(即不存在非一 X 的个案)。相比之下，拥有威权政体的个案集合却不是几乎全部人群。因此，集合 Y"包含"集合 X 的程度更大一些。如果推广这个观点，我们可以说，当存在多个必要条件时，出现频率较低的那些条件(即在相关的总体中是比较少见或异常的条件)就是比较重要的。

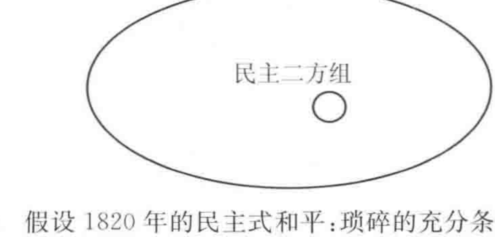


X_2：有地心引力的个案；
X_1：有威权政体的个案；
Y：社会革命的个案。

图 2.3 琐碎的必要条件：一个集合论例子

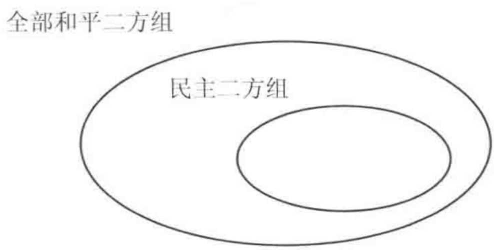
对充分条件（或者诸多因素组合后成为充分条件）的重要性的评估方式也类似。在这种条件下，*X* 是 *Y* 的子集。当 *X* 对 *Y* 的涵盖度增加时，充分条件 *X* 将变得更加重要；也就是说，当子集 *X* 趋于与 *Y* 有完美重叠时，*X* 就变得更加重要。因此，*X* 的重要性程度是变化的，它取决于结果的充分条件在多大程度上也接近于其必要条件。对于结果来说，非常重要的充分条件就是那些接近必要条件的条件。一个完全不重要的充分条件是这样的，即如果其存在就会产生结果，但是该条件永远不存在，因此永远不会产生结果。如果推广这个观点，我们可以说，对于既定的一类结果来说，当区分出多个充分条件的时候，出现频率多的那些条件就是比较重要的。

上文的民主式和平个案再一次提供了很好的阐释。如上所言，可以将民主式和平表述为“民主二方组和平的充分条件”。①在这个假设下，*X*（民主二方组集合）是 *Y*（和平二方组集合）的子集。然而，这个子集的相对规模随时间而变。我们假定图 2.4a 描述了 1820 年的情况。在那个时代，世界上只有几个民主政体，因此只有很少的民主二方组。*X* 集合（民主二方组）是和平二方组（即 *Y* 集合）的一个小子集。如果推进到 2000 年，如图 2.4b 所示，世界上一半以上的国家现在是民主政体。这就意味着，*X* 作为 *Y* 的子集已经急剧增加，构成了 *Y* 集合的相当大一部分。就 1820 年的数据而言，似乎可以声称民主式的和平理论是真实的，但它又是相当微不足道的发现。然而在当代，我们就不能说民主式和平是微不足道的了，因为民主二方组构成了所有和平二方组的一个相当大的子集。



a　假设 1820 年的民主式和平：琐碎的充分条件

① 实际上，如上所言，这种表述存在着潜在的问题，因为非战争不等于和平。



b 假设 2000 年的民主式和平：重要的充分条件

图 2.4 琐碎对重要的充分条件：民主式和平

2.7 结论

本章是数学导引，笔者勾勒了逻辑学和集合论的一些基本原则，因为它们与社会科学中的经验研究有关。我们的目标并不是简单地介绍逻辑学和集合论。相反，我们关注的来自逻辑学和集合论的观点都暗含地激活了定性研究领域。在此过程中，我们比较了逻辑学-集合论与概率论-统计学，开启了对两种传承的分析。我们已经看到在二者之间有很多类比；这些类比是有用的，同时也有误导性。它们一开始是有用的，但是当将它们看做另一方的完全对等物的时候，这种类比就具有了误导性。在逻辑学-集合论和概率论-统计学之间，存在一个转译问题。

本章所讨论的初始对比将在后续章节中再次出现。最后，这两种传承之间的数学差异在所有的研究领域（从研究目标、因果模型、概念和测量再到个案选择步骤）中都有广泛的分支。

参考文献与推荐读物

在哲学领域中有很多优秀的逻辑学导论性文献。对于社会科学家来说，可采用一部经典教材，即科皮、科恩和麦克马洪（Copi, Cohen and McMahon, 2010）。其他优秀的著述包括豪斯曼、卡亨和蒂德曼（Hausman, Kahane and Tidman, 2010），根斯勒（Gensler, 2002）和苏

佩斯(Suppes, 1999)。对模糊集进行全面数学介绍的著作是克利尔和袁(Klir and Yuan, 1995)。有关模糊集的开创性论文，亦可参见扎德(Zadeh, 1965)。本章讨论的许多观念都取材于拉金(Ragin, 1987, 2000, 2008)；也可参见施奈德和瓦格曼(Schneider and Wagemann, forthcoming)以及罗尔夫林(Rohlfing, forthcoming)。至于在社会科学中利用逻辑学和集合论进行研究的成果文献目录，参见COMPASSS 网站 http：www.compasss.org/。

Braumoeller, B. 2004. "Hypothesis Testing and Multiplicative Interaction Terms." *International Organization* 58:807—820.

Braumoeller, B., and G. Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44:844—858.

Cioffi-Rvilla, C. 1998. *Politics and Uncertainty: Theory, Models and Applications*. Cambridge: Cambridge University Press.

Copi, I.M., C. Cohen, and K. McMahon. 2010. *Introduction to Logic*, 14th edition. Upper Saddle River, NJ: Prentice Hall.

Downs, G. 1989. "The Rational Deterrence Debate." *World Politics* 41:225—237.

Eliason, S.R., and R.Stryker. 2009. "Goodness-of-fit Tests and Descriptive Measures in Fuzzy-set Analysis." *Sociological Methods and Research* 38:102—146.

Gensler, H.J. 2002. *Introduction to Logic*. London: Routledge.

Goertz, G. 2003a. *International Norms and Decision Making: A Punctuated Equilibrium Model*. New York: Rowman & Littlefield.

Goertz, G. 2003b. "The Substantive Importance of Necessary Condition Hypotheses." In *Necessary Conditions: Theory, Methodology, and Applications*, edited by G. Goertz and H.Starr. New York: Rowman & Littlefield.

Goertz, G. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41:

88—109.

Hausman, A., H.Kahane, and P.Tidman. 2010. *Logic and Philosophy: A Modern Introduction*, 11th edition. Boston: Wadsworth.

Kam, C., and R.Franzese. 2008. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis: A Brief Refresher and Some Practical Advice*. Ann Arbor: University of Michigan Press.

Klir, G.J., and B.Yuan. 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ: Prentice Hall.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.

Lakoff, G., and R.E.Núñez. 2000. *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics Into Being*. New York: Basic Books.

Mackie, J.L. 1965. "Causes and Conditions." *American Philosophical Quarterly* 2:245—264.

Mackie, J.L. 1980. *Cement of the Universe: A Study of Causation*. Oxford: Clarendon Press.

McNeill, D., and P.Freiberger. 1994. *Fuzzy Logic*. New York: Simon and Schuster.

Ragin, C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. *Chicago*: University of Chicago Press.

Rohlfing I.Forthcoming. *Case Studies and Causal Inference: An Integrative Framework*. New York: Palgrave Macmillan.

Russett, B. 1995. "The Democratic Peace: 'and yet it Moves.'"

International Security 19:164—175.

Schneider, C., and C. Wagemann, Forthcoming. *Set-theoretic Methods in the Social Sciences*. New York: Cambridge University Press.

Seawright, J. 2011. "Matching Quantitative Case Selection Procedures with Case-study Analytic Goals." Manuscript, NorthWestern University.

Smithson, M. 1988. "Fuzzy Set Theory and the Social Sciences: the Scope of Applications." *Fuzzy Sets and Systems* 26:1—21.

Suppes, P. 1999. *Introduction to Logic*. Mineola, NY: Dover.

Waltz, K. 1979. *Theory of International Relations*. Boston: Addison-Wesley.

Zadeh, L.A. 1965. "Fuzzy Sets." *Information and Control* 12: 338—353.

第一部分　因果模型与推断

3 先因后果与先果后因

原因(Cause):[中古英语,源自古法语,最终源自拉丁文动机(causa);理由;目的]名词,1.a.一个效果、结果或结局的生产者;b.为某个行动或结果承担责任者,如一个人、一个事件或一个条件。

——《美国传统英语词典》

3.1 导言

让我们区分两种不同的询问和解决因果问题的方式。一种方式始于一个结果(即 Y),然后回溯其诸多原因(即多个 X);第二种方式则反其向而行之,始于一个潜在的原因,然后追问其对 Y 的影响。前一种程序通常被称为"先果后因"路数(causes-of-effects approach),后者为"先因后果"路数(effects-of-causes approach)。追问全球变暖的原因,这是一个先果后因问题。探究炭排放对全球气温的影响就是一个先因后果问题。好的科学对这两种问题都关心。然而在社会科学中,沿着这两种研究路数会有重大的方法论后果,这里讨论的许多主题都是由这些差异引出的。

在多大程度上以及通过哪些方式探究先因后果和先果后因问题,定量和定性传承有差别。定量学者作为一个群体,明显倾向于将先果后因路数作为社会科学研究的标准方法。特别是,他们一直喜欢在总体或样本内估计某些特殊变量的平均效应。在这个传统下,学者将受

控实验看做学术研究的黄金标准。在实验中，人们试图估计一种处理的平均效应。类似地，在针对观测数据的统计研究中，人们试图估计所关注的某个自变量的平均效应。虽然定量研究者偶而会通过解释掉的方差最大化（例如，较高的 R^2）来解释 Y，但这种做法在当代定量研究中已经式微。在我们的调查中（参见附录），只有 6%的定量文章明确地讨论了统计量 R^2。

相比之下，在定性传承中，学者的兴趣点在于解释个案的结果，同时研究某些特殊要素在个案范围内的效应何在。这些学者通常始于在现实世界中发生的事件，并返回来询问其原因。与历史科学中（包括自然史、地质学和宇宙学）的其他学者类似，他们也提出先果后因模型，并用一些方法去确定过去发生的特殊事件的原因。理想地讲，诸多充分条件会带来结果，这些模型会确定这些条件（包括所有的重要条件）的组合。

对一个先果后因问题的回答几乎总是需要一个多变量解释。对一些特殊个案的结果的因果分析涉及大量的要素。相比之下，在统计研究的平均效应路数（average effect approach）下，就很容易并且通常只聚焦一个自变量。当然，在一种实验环境之外，研究者几乎总是需要在统计模型中加入其他"控制"变量来评估感兴趣的效应。但是，加入的这些其他变量是为了处理混淆原因问题，不是因为分析者的兴趣在于这些变量有效应或对所关注的结果有贡献。

定性学者也考察个体原因所带来的结果，这或者作为一种先果后因路数的一部分，或者是为了自身的目的。然而，他们不会将对原因的效应的分析等同于平均因果效应分析。相反，他们要进行这样的因果分析，即对于特定个案中的特殊结果来说，要询问诸多要素是否为必要条件或组合后是不是充分条件。这样的话，当定性研究者考察原因之结果时，他们潜在或显在地用那些与集合论和逻辑学一致的方式来概念化那些结果。

研究者在多大程度上对总体还是个案感兴趣，上述这些差异与此问题密切相关。根据定义，定量分析者几乎都集中关注总体分析，对个案不感兴趣。这自然促使他们依据应用于总体的平均值来考察原因的

结果。相比之下，尽管定性传统中的研究也涉及跨个案的推广，但是他们总是强烈地关注对个案的解释。这使得定性学者一点也不关注平均效应，却集中关注哪些原因带来了特殊个案中的结果。

3.2 研究问题的类型

某些问题具有这样的一般形式："什么引起了 Y?"定量和定性研究者都对解决此类问题感兴趣。然而，他们对这个问题的转译却不同。定量研究者用他们版本的先因后果路数(effects-of-causes approach)来转译这个问题："在由个案构成的一个总体中，X 对 Y 的平均效应是多少?"这些学者很少将该问题转译为他们版本的先果后因路数："对于一个总体来说，有哪些 X 可以解释 Y?"相比之下，定性研究者通常用他们的先果后因路数来转译该问题："对于一个或多个特殊个案来说，哪些 X 可以解释 Y?"定性研究者也可能用他们版本的先因后果路径来转译该问题："在一个或多个个案中，X 引起 Y 了吗?"

统计学方法论者认识到了先果后因与先果后因路数的差异。例如，统计学家霍兰论证道，在这两种路数之间存在一个"无法逾越的鸿沟"：

> 这两种路数都希望赋予短语"A 引起 B"以意义。(先果后因路径)通过将"A 引起 B"解释为"A 是 B 的一个原因"。(先因后果)模型却将"A 引起 B"解释为"A 的结果是 B"。(Holland, 1986:970)

虽然霍兰的引文强调了单个原因项，不过先果后因问题却使他们自己自然地设计出一种旨在"充分地"说明结果的多变量解释。按照先果后因路数，人们则要从结果 Y 开始，然后力图提出一个因果模型来确认能够解释 Y 的诸多条件。

尽管他们识别了这两种路数，当代定量研究者却坚持先因后果路数。正如莫顿和威廉斯所言：

> 政治科学的很多定量研究——我们可以说模型研究……聚焦于探索特殊个案的结果。有时候，所倡导的这种活动成为努力建立结果之原因的一般模型的一部分，但如果这样一个目标存在于研究者的头脑中，它通常就是隐含的。(Morton and Williams, 2010:35)

例如，当定量研究者询问，"是什么引起了民主"，他们通常探求感兴趣的特定自变量的结果。他们会问，诸如经济发展(Londregan and Poole, 1996)，政党(Mainwaring, 1993)和总统制(Cheibub, 2007)等变量是否对民主有某种影响。这些研究的核心恰恰是探索发展、政党或总统制在一个总体中是否对民主有影响。

作为定量研究基础的基本实验范式使这些分析者很难真正坚守先果后因传统。不重视先果后因研究是很自然的。例如，一些定量方法论者认为，先果后因研究只是"描述性的"(Sobel, 1995)或不能产生一般的因果知识(Beck, 2006)。根据安格里斯特和皮施克(Angrist and Pischke, 2009:5)的观点，"不能用任何实验来回答的研究问题都是基本上不能确认的问题"(fundamentally unidentified questions, FUQ)。分析者指出，影响深远的 Neyman-Rubin-Holland 因果模型"是一个纯先因后果模型。至于我们如何从结果集合走向先果后因模型，该先因后果模型根本不关心"(Morton and Williams, 2010:99)。当统计学家戴维(Dawid, 2000)提出将先果后因路数作为一个特殊的因果性的时候，他没有引起重视。在回应来自很多杰出统计学家的一系列评论时，他对自己提出的先因后果分析没有激起什么讨论表示吃惊。"我建议进行'先果后因'推理。原本期望它最引起争议，我惊讶的是，目前的讨论与我的建议很少有关联(Dawid, 2000:446)。

有人可能会问，统计研究者会怎样解决先果后因问题。在追寻"解释 Y"的目标时，一种途径是尽力使解释的方差最大。在早期的统计学

实践中，解释的方差（即令 R^2 最大）是研究的一个主要目标，也是评价统计模型的关键标准。在这条路数中，如果模型的 R^2 接近于 1.00，它就解释了 Y。有来自几十年前的许多统计学文章可作为例子，那时候研究者感兴趣的问题有：(1) 每个 X_i 解释 Y 的方差是多少，以及 (2) 整个模型解释 Y 的方差有多少。

然而在政治科学和社会学领域，大约始于 20 世纪 80 年代中期，统计学者开始拒绝“R^2 模型”，转而接受先因后果路数。领军的定量方法论者开始相当怀疑解释方差的目标何在：

> 如果你的目标是得到一个大的 R^2，那么你的目标就不同于回归分析设计的目标……最好的回归模型的 R^2 通常低于其他方式获得的 R^2。（生成）一个大的 R^2 的目标……和任何政治科学研究问题不大相关。(King, 1986:677；也参见 King, 1991)

金和其他学者表明，将接近于因变量的某些变量（例如，Y 的滞后值）加入模型后，很容易使 R^2 增大。而且，寻找大的 R^2 值会鼓动加入许多自变量，并且由于多种原因，这种做法可能是有问题的（Achen, 2005）。总之，现在很少有定量研究者用 R^2 统计量作为评估因果模型的基础。在目前发表的显示度高的著述中，R^2 完全可能小于 0.10。

还值得注意的是，在尝试或多或少综合地解释结果的过程中，早期的统计学者经常提出定量的通径模型（path models）（如 Blalock, 1964）。这些学者利用诸如结构方程模型之类的技术，来设定位于一个序列中不同点的多个自变量如何共同作用才能产生所关注的结果（Bollen, 1989）。在社会科学之外的统计分析工作中，虽然定量的通径模型仍然会偶尔被探讨（例如，Pearl, 2000; Morgan and Winship, 2007），但是在社会科学的实际实践中，它们基本上已经被经验研究放弃。

相比之下，定性研究者也提出因果论证，他们意在设定一些要素，这些要素组合起来作为充分条件会带来结果。例如，定性研究者试图确认第一次世界大战、东亚的异常发展、冷战的结束、异常慷慨的福利

国家的产生以及新民粹主义政权(neopopulist regimes)上升等现象的原因。这种研究的核心目的是对研究范围内的每个个案的特殊结果都给出全面的解释(如 Levitsky and Way, 2010)。

定性学者也研究并分析个体原因有哪些结果。他们这样分析的同时，通常也试图全面地解释一个结果。要想聚焦于解释 Y(即分析一个结果有哪些原因)，就需要展示各种 X 是如何带来因果效应的(即分析诸多原因有哪些结果)。有时候，定性研究者对某个特定的原因及其效应本身感兴趣。例如，科利尔和科利尔(Collier and Collier, 1991)的《塑造政界》(*Shaping the Political Arena*)就集中阐释了，在拉丁美洲，劳工整合期(labor incorporation periods)* 的种种变化如何影响到长期的政治动向。分析的主要目标是展示劳工整合期的各种异同点为什么对于解释个案**之间的主要政治异同点来说是根本。

然而，当定性研究者用先因后果路数时，即使当他们向一个总体推广时，也不估计平均的因果效应。相反，他们通常会研究的是，可以将哪些条件带来的效应理解成在一个或多个个案中实际发生的某个结果的必要条件。因此，当定性研究者断言 X 对 Y 产生了一个因果效应的时候，他或她通常相信，如果 X 没有发生(或以另外的方式发生)，Y 就不会在所研究的特定个案中发生(或发生的方式不同)。①例如，科利尔和科利尔(Collier and Collier, 1991)的研究建立在如下观念之上，即不

* 科利尔和科利尔的《塑造政界》(*Shaping the Political Arena*)一书比较了 8 个拉美国家(阿根廷、巴西、智利、哥伦比亚、墨西哥、秘鲁、乌拉圭和委内瑞拉)的政治系统在 20 世纪上半叶的形塑与发展。作者们展示了国家和政党是如何对有组织的工人阶级的兴起做出反应的，这种反应对于形塑政治联盟、政党系统的稳定或冲突的模式、政权及其变迁的轮廓都至关重要。本书关注这些国家如何控制和动员涌现出的劳工运动(不同的控制和动员策略产生了不同的冲突和适应模式)，探讨国家在塑造劳工运动时的策略，解释不同渠道的国家政治变迁与长期的城市、商业和工业发展历史。本书是关于拉美政治领域的一部杰作。在书中，"整合"(incorporation)一词的含义因国家不同而有别，"整合期"(incorporation periods)在不同国家也不一样。作者们区分了四类广义类型的整合期(即四对国家)，每个整合期都有一个共同的特征：国家努力控制工会。——译者注

** 即国家。——译者注

① 当某个特定个案中的结果被多种充分条件过度地决定了的时候，这便是例外情况。在这种情况下，个体要素 X 在该特殊个案中可能不是 Y 的充分条件。

同类型的劳工整合期会遗留各自不同的独特遗产。尽管劳工整合期不完全决定政治后果,但对于在所研究的个体国家中实际发生的许多重要的政治动向来说,它们却是必要的组成部分。

3.3 个案

定性研究者受吸引去关注有关某些重要结果(如民主、战争、经济增长和制度变迁)之原因的一些普遍问题。同时,针对个案中的这些结果,他们也努力给出解释。那种先果后因路数导致他们给出的解释同时适用于一组个案和该组内的每个个案。在定性传承中,提供令人信服的普遍解释的同时也对个案提供了令人信服的解释。

因此,定性学者坚持以下基本原则:

> *对 Y 的一种好的普遍解释也是对 Y 中个案的一个好的解释。*

例如,为了给欧洲两次世界大战之间的社会民主一个好的定性解释,需要确认在瑞典、挪威和丹麦的社会民主的诸多原因。在吕贝特(Luebbert, 1991)的著名研究中,相同的一些要素(分裂的中产阶级、弱自由主义、社会主义者和农民党的联盟)可以在两次世界大战之间的欧洲引起普遍的社会民主,也在瑞典、挪威和丹麦这些特殊个案中引起社会民主。对某个结果的一般性解释如果不能很好地解释个案,那么该解释就不是一个好的一般性解释。再举另一个流行的例子,在历史上没有受制于殖民主义的农业国家中为什么出现社会革命,斯考切波(Skocpol, 1979)提出了一个解释。她的解释是一般性的,因为该解释意在应用于此(限定的)范围内的全部社会革命实例。同时,设计这种解释的目的也是为了充分地说明在其范围内的每一个实例,即法国、俄罗斯和中国。如果该解释不能阐明每一个个案中的社会革命,就不能认为它是一个成功的全面解释。

如此看来，定性学者需要确保其因果模型在其个案中也有效。他们希望确证，因果模型（在通常适度的范围内）可以解释所探究的全部（或几乎全部）个案中的 Y。这样的话，他们不会将评估一个显著的平均结果作为目标，但他们会努力给出一个因果论证，使该论证几乎总是可以解释其 $Y=1$ 的个案。

相比之下，在定量的先因后果分析中，对个案及其原因进行集中研究至多扮演一种次要角色。为何如此？如果考察个体观测项在一个典型实验——如著名的米尔格拉姆（Milgram，1947）顺从实验——中的位置，即可理解其原由。通过操纵实验条件（例如，与“老师”的距离远近），个体便被随机地分配到实验组和控制组。空间的接近如何影响顺从方面的权威人格，虽然实验可以告诉我们关于这方面的重要内容，但是，实验的设计并不是为了解释为什么任意一个特定的被试按照自己的方式行事。实验告诉我们的是一般意义上的结果，不会告诉我们某些特殊个体的结果有哪些原因。在一次实验中，甚至都不存在一种手段去争论个体个案，因为更大的研究社群可能都不了解个体的身份。

同样，在定量研究中，关注点不是去解释任何一个个案。毋宁说，关注点是某个自变量在总体中的平均效应。这个平均值可能或不能应用于特殊个案。至于一个自变量是否会在某个特殊观测项上施加其通常的效应，这个问题不在考虑范围之内。人们可以不知道 X 在任何单独个案中如何影响 Y，却可以知道 X 在总体中对 Y 的通常效应。

原则上讲，人们可以在一个大样本（larger-N）先因后果分析中询问个体个案：对于处于实验组的个体 i 来说，X 有一个因果效应吗？然而，如果没有任何来自个案内部的信息，就很难回答这个实践问题。实际上，标准的统计学路数有一个假定，即不可能估计个体 i 的因果效应，这恰恰是为什么要针对一个总体来估计平均因果效应（如 King，Keohane and Verba，1994）的原因。

的确，关于平均因果效应的信息可应用于个体观测。例如，医生会例行公事地基于大样本的实验和/或观测研究向个体提出建议。如果他们开出了药方并且某些症状消失，就可能推断药物对病人个体产生了一个因果效应（当然，它也可能是安慰剂效应）。然而，如果不分析药

物影响症状的因果途径,这种推论就只是信息充分的猜测,很难是其他信息。

3.4 结论:互补性与混合方法研究

是接受先果后因路数,还是先因后果路数,这在很大程度上要围绕着对比性的研究导向和目标来决定:是通过个案内分析来研究个案,还是通过个案间分析去考察总体中的集中趋势。定性学者用个案内分析来做出关于产生结果的特殊事件和进程的推断。个案内分析密切协作,在所研究的特定个案中,努力解说哪些因素会产生结果。相比之下,定量学者用个案间分析做出关于总体的推断。在这里,人们自然会追问并回答这样的问题,即所关注的特殊变量在作为整体的总体内会产生哪些典型的效应。但是,要是论说某个变量对任何一个特殊个案有什么效应,这并不常见。

每一种路数都能并且的确能够相互补充,这个事实表明了二者的功用。当定性研究者努力解释他们的特殊结果的时候,他们通常始于既存的知识,包括先前关于特殊原因的结果的统计研究。与之类似,很多研究探讨的是在特殊个案中的结果有哪些原因,好的定量研究通常要考虑此类研究的结论。这些定性的发现应在统计分析中经受进一步的评估。

如此看来,某些有力的理由让人相信定性与定量路数都是有价值的,并且可以互补。对于一个小样本中的某个结果来说,一个好的定性解释会促使人们考量:如果广义地理解范围,促成一个大样本分析(其目标不再是解释特殊个案,而更多是估计平均结果),那么同样的一些要素是否还起作用。同样地,当报告关于先因后果的定量结果时,自然会询问:这些结果是否可以从个案的历史角度来理解;人们希望尝试在特殊个案中锁定这些效应。这些类型的互补性使混合方法的研究成为可能,并且它们指明了跨界的交流与合作的价值。

参考文献与推荐读物

霍兰（Holland，1986），戴维（Dawid，2000），莫顿和威廉斯（Morton and Williams，2010）讨论了先果后因路数与先因后果路数的差异。在（如实际践行的）政治科学和社会学研究中，尽管这些路数与定性及定量研究范式息息相关，但它们不需要总是如此。至于实验和统计研究通过什么方式有助于检验大样本的先果后因模型，有关这方面的讨论参见赫克曼（Heckman，2005），莫顿和威廉斯（Morton and Williams，2010：33—41，第6章）。至于定性研究通过哪些方式有助于估计平均效应，这方面的讨论参见耶林（Gerring，2007）。有关个案在定性研究中的角色，参见科利尔、布雷迪和西赖特（Collier，Brady and Seawright，2010），乔治和贝内特（George and Bennett，2005），马奥尼和吕申特梅耶（Mahoney and Rueschemeyer，2003）。

Achen, C.H. 2005. "Let's Put Garbage-can Regressions and Garbage-can Probits Where They Belong." *Conflic Management and Peace Science* 22:327—339.

Angrist, J.D., and J.-S.Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Beck, N. 2006. "Is causal-process observation an oxymoron?" *Political Analysis* 14:347—352.

Blalock, H.M., Jr. 1964. *Causal inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

Bollen, K.A. 1989. *Structural Equations With Latent Variables*. New York: Wiley.

Cheibub, J. A. 2007. *Presidentialism, Parliamentarism, and Democracy*. Cambridge: Cambridge University Press.

Collier, D., H.E.Brady, and J.Seawright. 2010. "Sources of Leverage in Causal Inference: toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E.Brady and D. Collier. Lanham, MD:

Rowman & Littlefield.

Collier, R.B., and D.Collier. 1991. *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America*. Princeton: Princeton University Press.

Dawid, A. P. 2000. "Causal Inference without Counterfactuals (with Discussion)." *Journal of the American Statistical Association* 95:407—448, 450.

George, A., and A. Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Gerring, J. 2007. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Heckman, J.J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35:1—97.

Holland, P. 1986. "Statistics and Causal Inference(with Discussion)." *Journal of the American Statistical Association* 81:945—960.

King, G. 1986. "How not to Lie with Statistics: avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30:666—687.

King, G. 1991. "'Truth' is Stranger than Prediction, more Questionable than Causal Inference." *American Journal of Political Science* 35:1047—1053.

King, G., R.Keohane, and S.Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Levitsky, S., and L.A.Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. NewYork: Cambridge University press.

Londregan, J., and K.Poole. 1996. "Does High Income Promote Democracy?" *World Politics* 49:1—31.

Luebbert, G. 1991. *Liberalism, Fascism, or Social Democracy:*

Social Classes and the Political Origins of Regimes in Interwar Europe. Oxford: Oxford University Press.

Mahoney, J., and D. Rueschemeyer, eds. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

Mainwaring, S. 1993. "Presidentialism, Multipartism, and Democracy: the Difficult Combination." *Comparative Political Studies* 26:198—228.

Milgram, S. 1974. *Obedience: An Experimental view*. New York: Harper & Row.

Morgan, S., and C. Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Morton, R., and K. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Skocpol, T. 1979. *Statesand Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.

Sobel, M. 1995. "Causal Inference in the Social and Behavioral Sciences." In *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, edited by G. Arminger. New York: Plenum Press.

4 因果模型

因果模型必然有不同的形式，这取决于用它们来做什么以及研究的是哪类系统。

——南希·卡特赖特(Nancy Cartwright)

4.1 导言

本章将比较两个模型。一个是加法的线性因果模型，它常见于统计学传承中；另一个是基于逻辑学的集合论因果模型，它常(隐含地)用于定性研究。在后续几章中，我们将探究一个或另一个因果模型的分支，因为它们暗设着关于因果性、因果不对称性、反事实等的不同观点。

我们首先通过例子来介绍和展示这两种因果模型。接下来的几节将描述二者的实质区别。我们认为，虽然这两个因果模型有显著差异，但哪一个都没有先天的正确性。它们都可以潜在地用来解释社会和政治现象。然而，由于二者有别，在将它们结合或利用来自一种模型的洞见去推进另一模型的道路上，还是障碍重重。因此，我们建议，为了克服这些障碍，第一步就是领会二者各有什么独特之处。

4.2 两种因果模型

在定性和定量研究中使用的标准因果模型很相似，只在一些不明显的方面有差异。为了说明这一点，现在考虑下面两个模型，它们分别在定量和定性范式中很常见：

$$Y = \beta_0 + \beta_1 X_1 + \gamma_1 C_1 + \gamma_2 C_2 + \cdots + \varepsilon \tag{4.1}$$

$$(A\text{ 与 }B\text{ 与非 }C)\text{或}(A\text{ 与 }C\text{ 与 }D\text{ 与 }E)\text{对 }Y\text{ 来说是充分的} \tag{4.2}$$

方程(4.1)的典型体现包括像 OLS 回归这样最流行的一般模型，以及如双重差分回归(difference-in-differences regression)这样正在流行的其他模型。方程的基本形式也包含对数线性模型(log-linear model)和多项式回归(polynomial regression)。虽然 logit 模型采用非线性函数，但它们在指数上是线性的。logit 模型之所以经常被使用，是因为它们可以处理二分因变量，并不是因为分析者相信因果模式是非线性的。①分析顶尖杂志上发表的研究可以发现，在定量传统中建立因果模型时，方程(4.1)家族的某个成员是最常使用的。

就我们的目的而言，方程(4.1)的一个关键特征是，研究者的兴趣点是估计 X_1 对 Y 的效应(即 β_1)。除了与核心假设相关联的变量(即 X_1)之外，该模型还包括几个控制变量或混淆变量(即 C_i)。仅关注 X_1 和 Y 之间的一个显著的双变量效应 $\hat{\beta}_1$ 鲜有说服力；人们必须对这样一种关注点做出回应，即在控制变量 C_i 后，该效应可能消失。在方法论文献中，人们对方程(4.1)做出了大量的改进，但是在大多数利用定量方法的高显示度的实际研究中，这种基本的、加法线性函数形式仍然是常规的做法。

---

① 用来估计非线性模型的统计方法的确存在，例如贝茨和沃茨(Bates and Watts, 1988)，但是在政治科学和社会学中通常不讲授这些模型。

如果转向定量方法论者之间关于最佳实践的争论，那么诸多子传承(subculture)之间的差异就变得明显。最值得注意的是，与一般线性模型相比，Neyman-Rubin-Holland 模型或"潜在结果"架构("potential outcome" framework)提出了一种不同的统计研究路数，这种路数与理想化的随机实验有相当清楚的紧密联系(参见 Morgan and Winship, 2007)。这一架构的基础模型是：

$$因果效应 = \bar{Y}_{(X=1)} - \bar{Y}_{(X=0)} \tag{4.3}$$

像方程(4.1)中的 Y 一样，这个"因果效应"也是一个随机变量，人们通常检验的假设是：该因果效应显著地有别于 0。这个架构比一般线性模型方法更加明确，后者的目标是估计 X 的平均因果效应。在这个基础模型中不需要控制变量，因为对处理的安排是在大量的观测中随机进行的。

虽然潜在结果架构在统计学方法论者中间相当有影响，但它在研究实践中还没有得到广泛应用(见附录)。①关键原因在于，当从二分变量转向连续变量，并引入其他细微项(如交互项或中介变量等)的时候，该架构就变得相当复杂。不管怎样，许多方法论者还会坚持认为，定量范式的基本因果模型应该是方程(4.3)，而不是方程(4.1)，尽管后者主导着研究实践。

在当代统计方法论者中间，贝叶斯方法(Bayesian approaches)是另外一种重要的子传承。虽然贝叶斯分析法已经出现了几十年，但是从历史上讲，它并没有被大量地运用在社会科学研究中。然而，随着近年来计算能力的提升与马尔可夫链蒙特卡罗方法(Markov Chain Monte Carlo methods)的提出，该方法在应用分析中的主要障碍已经减少。在政治科学领域中，定量方法论者中间还存在着一种生机勃勃的贝叶斯子传承(vibrant Bayesian subculture)(例如，参见 Schrodt, 2010 给出的热情支持的观点)。这种子传承始于一组相当不同于方程(4.1)的假定和原理(相关的精彩讨论，参见 Jackman, 2009)。可能由

① 珀尔(Pearl, 2000)针对因果性提出了另一个有影响的概率论方法，该方法在研究实践中也没有重要影响。

于贝叶斯子传承与频次学派（frequentist approaches）差异太大，因而不会（至少直接地）关心潜在结果架构。①

简而言之，在统计方法论学者中间有各种重要的小群体，他们内部争辩不休。然而，鉴于我们正尝试描述一些主导的研究实践，这些实践可以生产出被大的学术共同体看成是社会科学中的最佳学术研究，因此，方程（4.1）的基础模型显然最符合我们的目的。

在定性范式方面，我们建议将方程（4.2）作为众多研究的基础。这项建议不是没有争论的，因为定性研究者通常没有系统性，也不写出什么形式方程。然而，我们发现他们的论证经常隐含地采取一种与方程（4.2）相类似的形式。在这里，集合论思想代替了加法模型的代数和统计学。

前文用语言表述了模型，人们也可以用数理逻辑来表述模型。方程（4.4）的因果模型区分了生成某种结果的诸要素的两种组合，即这两种组合成为该结果的充分条件。这个模型使用了逻辑符号和布尔（Boolean）运算符。符号 $\wedge$ 代表着逻辑“与”，$\vee$ 代表着逻辑“或”，$\neg$ 意味着逻辑“非”，$\rightarrow$ 表示“充分性”：

$$(A \wedge B \wedge \neg C) \vee (A \wedge C \wedge D \wedge E) \rightarrow Y \qquad (4.4)$$

我们可以用集合论符号给出相同的方程，其中 $\cap$ 表示交集，$\cup$ 表示并集，$\bar{C}$ 表示集合 C 的补集，$\subseteq$ 表示子集或对等集：

$$(A \cap B \cap \bar{C}) \cup (A \cap C \cap D \cap E) \subseteq Y \qquad (4.5)$$

正如方程（4.4）和方程（4.5）所表示的，这些因果模型看起来很不同于方程（4.1）的定量模型。然而，换一种方式写同样的方程，会让它看起来更像其对应的定量模型：

$$Y = (A * B * c) + (A * C * D * E) \qquad (4.6)$$

当这样表达的时候，对集合论/逻辑模型的一个可能的反应便是声

① 例如，杰克曼（Jackman，2009）的一本有影响的著作中并没有“潜在结果”这个索引条目，也几乎没有参考因果效应方面的思想。

称这是一种谈论交互项的方式。但是正如在两种语言之间的转译那样，逻辑“与”和统计模型中的乘法仅在局部上可比拟。如我们在“数学导引”一章所讨论的那样，还有相当多的差异是这种比拟所不能把握的。

我们也可以用方程(4.6)将定性模型转译成统计学术语：

$$Y = \beta_1(A * B * c) + \beta_2(A * C * D * E) + \varepsilon \qquad (4.7)$$

谁会在一篇研究文章中见过这样的模型？答案是：没有人。没有什么可以阻止统计软件去估计这样一个模型。然而，标准的统计学建议却提供了很好的理由去拒绝这种做法。例如，这里没有截距项。要想估计方程(4.7)，需要假定截距是0，而它可能就是我们要检验的假定，不能仅假定它就罢了。相比之下，“截距”概念在集合论/逻辑学范式中意义不大。

4.3 集合论模型和个案内分析：一种阐释

大多数社会统计学者都在一门或多门统计学课程中学习过一般线性模型。然而，集合论的因果模型对于许多人(通常是那些隐含地使用它们的定性研究者)来说是不熟悉的。因此，考虑一下像方程(4.5)这样的式子在实践中采取什么形式是有帮助的。为了便于说明，我们考察布雷迪(Brady, 2010)所分析的媒体提前宣布(early media call)的影响：在2000年的总统大选中，媒体提前宣布了阿尔・戈尔在佛罗里达州的胜利。这个例子会表明一个集合论模型是怎样隐含地用于对某个特定结果的个案内分析的。

布雷迪的研究向洛特(Lott, 2000)有影响力的结论提出了质疑，后者断言，在佛罗里达州的10个走廊县(panhandle counties)，乔治・W.布什至少丧失了1万张选票。在这些县里，当大选投票还没有结束的时候，网络就过早宣布了阿尔・戈尔获胜。布雷迪声称，洛特不恰当

地使用了双重差分回归模型(参见 Angrist and Pischke，2009)，做出了错误的推断。

布雷迪没有使用个案间统计方法论，而是利用个案内分析来估算丧失的选票数量。为什么媒体过早宣布会造成布什的选票丧失，否则他就会得到这些选票呢？布雷迪考察了其中必然存在的机制。需要特别指出的是，布雷迪开展了一系列"箍筛检验"(hoop tests)(有关这种检验的讨论，参见第 7 章)，他在检验中区分出一些充分条件，即哪些条件使得媒体提前宣布导致布什失去了佛罗里达居民的选票。

布雷迪的箍筛检验指出，布什要是丧失选民 i 的选票，选民必须同时是三个集合的成员。首先，必须是居住在佛罗里达州东部走廊县的居民集合；其次，没有投过票的合法选民集合；再次，听到了媒体声称阿尔·戈尔获胜的选民集合。用逻辑学术语讲，每个集合中的成员都足以使得该个体被编码成不给布什投票的潜在选民。如此看来，他(隐含地)使用了下面的集合论方程：

$$Y \subseteq L \cap E \cap H \tag{4.8}$$

在这个方程中，L 代表东部走廊县的居民；E 是在媒体没有公布前投票的合格选民；H 选民听过媒体公布戈尔当选；Y 是由于早公布戈尔当选而使得布什最多可能丧失的选票数。这个方程说的是，可能改选者的集合是 L、E 和 H 三个集合的交集。该方程表示的是一种简单但有用的、计算最多可能出现的选票损失数的方法，即在这三个集合的交集中的成员总数。

然后，在布雷迪的工作中，一个关键的挑战是：如何估算由处于所有这三个集合中的成员构成的总体的规模。他利用佛罗里达州的先前投票模式、媒体曝光和通常的投票行为知识，通过个案内分析来估算。根据布雷迪的估算，"布什丧失的选票数上限约为 224，……实际丧失的选票数大约在 28 至 56 票之间。洛特给出的 1 万票根本就讲不通"(Brady，2010：240)。

这个简单的例子说明，在集合论因果模型和一般线性统计模型之间有三点关键差异：

(1) 无控制变量。例如,在双重差分设计中,针对佛罗里达的每个走廊县都有各种控制变量(亦称固定效应)。

(2) 因变量不同。方程(4.8)给出了最大可能的选票丧失估计量。正如我们在第 2 章所讨论的那样,在集合论中,最大(或最小)值经常在统计模型中扮演均值的角色。

(3) 必要和充分条件。布雷迪的模型假定了一系列条件,其中的每一个都是可能丧失选票的充分条件。

出于我们考虑的目的,问题并不是洛特的统计模型与布雷迪的集合论模型孰优孰劣,而是他们的工作完全不同,考虑不同的因变量,并产生不同类型的结果。

4.4 因果关系的复杂性

定量和定性研究者都假定,现实世界中的因果模式在某些方面确实是非常复杂的。他们同样也认为,对因果关系进行建模本身就是一件复杂的工作。然而,因果复杂性的形式因定量和定性范式而不同。

在定量传承中,分析者要假定一个因变量总是存在着多个原因,这里体现了因果复杂性。例如,方程(4.1)就包含了许多自变量。此外,误差项(即 ε)经常被解释为是部分由缺失的自变量构成的。在这个传统中,人们经常假定原因太多,以至于不可能全部找出来。

虽然因变量的原因很多,但是定量研究的焦点经常是特定的一个自变量。例如,在方程(4.1)表示的因果模型中,关注点是 X_1;其他自变量被看做"控制变量",并且研究者对它们不是特别感兴趣。因果复杂性提出的挑战是,在已知其他变量存在的情况下,如何提出一种好的估计量来估计感兴趣的变量的平均效应。这个挑战是令人苦恼的,因为某些其他原因会影响感兴趣的主要自变量和因变量(可引起潜在的虚假关系),因此,需要将它们区分出来并加入到统计模型中。更一般

地说，我们需要控制变量以便产生因果的同质组。①

相比之下，在定性范式中，因果复杂性体现在如下事实之中，即研究者建构的因果模式经常关注的是个案的组合（combinations）（Ragin，1987）。正如方程(4.2)所展示的那样，定性研究者常常要寻找能够产生结果（或作为其充分条件）的一些因果程序包或配方（causal packages or recipes）。我们可以在方程(4.7)中看到这一点，该方程用统计学术语重写了定性模型：各个 β 与因果程序包（而不是个别变量）有关。我们在布雷迪的例子中也可以看到这一点，在该例子里，三个集合的组合产生了丧失选票的最大可能值。

在寻找因果配方的过程中，单个变量的角色经常被轻视，尤其是当单个变量不是必要条件的时候。"原因 C 的平均效应是什么？"如果 C 的角色在不同的因果构型中变化，那么该问题可能就意义不大。例如，在方程(4.2)中，C 有时有积极影响，有时有消极影响，这取决于与之相伴出现的其他原因项。同样，B 在和 A 共同出现以及在 C 不出现时就重要，但是 B 在其他情形下就对结果没有影响。因此，要想充分地设定 B 的效应，需要说明 B 出现在什么情景（等其他原因项）中。

现在，在拥有统计交互项的情形中也会出现同样之事。在方程(4.1)中，我们呈现了最常见的统计模型，但是关于交互项的假设和模型不是不常见的（参见附录）：

$$Y = \beta_0 + \beta_1 X_1 + \beta_1 X_2 + \beta_3 X_1 * X_2 + \gamma_1 C_1 + \gamma_2 C_2 + \cdots + \varepsilon \tag{4.9}$$

在这个模型中，在 X_2 的一定范围内，X_1 的效应可能为正，而在 X_2 的另一个范围时，X_1 的效应可能为负。定性和定量传承都赞同的是，当存在交互项时，构成交互项的单个变量的效应是多少，人们都所言甚少。

然而，优秀的统计学家几乎从不估计如方程(4.6)那样复杂的集合论模型。为了估计该方程，统计学实务建议在模型中将低级项（如 A、

---

① 在有关因果性的哲学-统计学文献中（如 Cartwright，1989：55—56），这一点是关键。

AB、*AC* 和 *AD*)加入模型之中。虽然这种做法有非常好的统计原因，但是这些原因在集合论模型中是不存在的。我们可以假想提出一些统计方法，用它们对定性模型中给出的各类因果构型进行建模。但是，这些统计模型，如布尔概率单位(Boolean probit)和布尔 logit(Boolean logit)方法(Braumoeller, 2003)都在统计学主流之外。

事实上，集合论模型的因果表达根本不是交互项。它们是特殊的因果组合。在方程(4.6)中，我们没有 $X_1 * X_2$ 这类表达，但有一个涉及 *A*、*B* 及非 *C*(即 *c*)成员资格的特殊的因果关系组合。正是在这种特殊的 *A*、*B* 和 *c* 构型中的成员资格才是 *Y* 的充分条件。因此，在集合论模型中，逻辑操作“与”(AND)将一些原因项加在一起成为一个“程序包”，或可称为“充分性组合”(sufficiency combinations)(目的是突出这样的想法，即该组合对于结果来说是充分的)。这激励人们更多地考量整个程序包，而不是其单个组成部分(Ragin, 1987)。

在诸如比较历史分析之类的定性领域中，因果论证通常涉及这些充分性组合。这样的话，斯考切波(Skocpol, 1979)提出，国家崩溃和农民起义组合在一起，足以在一个农业-官僚社会中造成社会革命；唐宁(Downing, 1992)认为，中世纪的宪政与军事革命的缺乏也足以造成早期的民主；马奥尼(Mahoney, 2010)也认为，少量本地人口和高度经济发展是西班牙殖民早期的美国(early colonial Spanish America)高度社会发展的充分条件。在某些情况下，一个因果组合中的个体因素位于不同的时间点。例如，在吕贝特(Luebbert)关于两次世界大战之间的欧洲政权的论证中，第一次世界大战前的软弱自由主义者和第一次世界大战后的红-绿联盟(red-green alliance)都是产生社会民主的因果程序包的一部分。

4.5 殊途同归

在定性和定量范式中使用的因果模型之间还有另一个差异，该差

异是围绕“殊途同归”(equifinality)(George and Bennett，2005)或拉金(Ragin，1987)所谓的“多重因果性”(multiple causation)而生的。殊途同归指的是这样的观念，即有多重因果路数指向同一结果。在集合论的因果模型中，殊途同归可以用逻辑“或”(OR)来表示。例如，在方程(4.6)中有两条因果路径 ABc 或 $ACDE$；其中任何一条对于达到结果来说都是充分的。值得注意的是，殊途同归并不要求有原因结合(causal conjunction)。例如，下面的模型就表达了殊途同归而无须原因结合：$A \vee B \vee C \rightarrow Y$。殊途同归的独特之处在于，多个条件或条件的组合[而不是如前文讨论的结合因果性(conjunctual causation)]的存在能产生相同的结果。①在实践中，设计的定性因果模型所容纳的个案可以不在少数，通常既包含诸多原因项的结合，也包括殊途同归性，如方程(4.6)所示。

殊途同归的存在对于定性因果模型来说不是独一无二的。在诸如方程(4.1)这样的统计模型背后隐含的观念是，达到一个特定结果的潜在路径有成千上万条。统计方程的右侧本质上代表的是一个加权总和，并且只要该加权总和大于设定的临界值(例如，在 logit 分析中)，就可以预测结果会发生。运用这种殊途同归来解释方程(4.1)，就有无穷多种方法可达到因变量的某个特殊值。殊途同归肯定存在。的确，殊途同归普遍存在，以至于我们谈论它几乎没有什么意思了。

在定性研究范式中，只有少数因果路数能达到一个特定的结果，这个事实才使得殊途同归性成为一个有用的概念。每一条路数通常都是诸多因素的一种特殊结合，但是其数量不会太多。在范围条件通常有限下的定性研究中(参见第 16 章)，其目标是区分出在总体中存在的全部因果路径。

事实上，在定性研究中，分析者通常会尝试对所研究的全部个案进行分配，以便设定具体的因果路径。由于总的研究目标是解释个案，人

① 我们认为，关于殊途同归的大部分讨论都认为其特殊之处在于表达了多因素的结合，这是不恰当的。如金、基欧汉和维巴(King，Keohane and Verba，1994：87—89)那样，如果用一种统计学的视角来重点关注这个方面，那么殊途同归性无非就是一种讨论交互项的方式罢了。

们为此目标要区分出每个个案都遵循的因果路径。例如，希克斯等人(Hicks et al., 1995)的结论认为，有三个独立的路径刻意指向早期的福利国家，他们的分析可让我们确定到底有哪些个案遵循其中的每一条路径(也可参见 Esping-Andersen, 1990)。在定性研究中，这些因果路径在一般理论知识方面可以扮演非常重要的组织角色。另引一例，穆尔(Moore, 1966)的著名研究确认了三条指向现代世界的不同路径，每一条路径都通过诸多变量的某种特殊结合来界定，并且清楚地区分出遵循每一条路径的特定国家。

在定量研究中，根据基于自变量建立的通用因果构型来对个案分组，这种做法似乎也没有用。虽然人们可以做到一点，但它在该传统中不是惯常的做法。为了理解它为什么无用，有益的办法是：用逻辑学眼光来看统计方程(4.1)会像什么样子。在布尔代数(Boolean algebra)中，"+"号代表逻辑"或"，并暗示不同的原因或其组合，每一个都是结果的充分条件。因此，从逻辑学的观点来看，可以这样解读统计方程(4.1)，即它暗示了每一个变量本身对于 Y 来说都是充分的。大多数研究者都会认为这个命题是荒谬的，因为单个因素自身对于结果来说几乎从来不是充分的；只有因素的结合才是充分的。如此看来，正如方程(4.2)那样转译统计模型是没有意义的一样，将代数-统计模型转译为逻辑学也有问题。

事实上，统计方程(4.1)中的每一个变量都只是影响结果的众多潜在因素之一。现实是，整个模型就是该路径。关于这一点，有一个很好的展示来自多元方法研究(multimethod work)，该研究首先检验一个统计模型，然后用其结果来选择个案研究做深入分析(如 Lieberman, 2005)。在选择个案时，人们可能选择在线上或线外(或两者都有)的观测点，这取决于研究目标。但是，对于整个总体来说线本身就是一个独特的因果路径。完整的因果模型适用于全部个案，并作为对所有个案的一类解释。

在统计学传承中，从这种讨论中可引出关于使用殊途同归的一个关键结果。将统计模型看成是假定了殊途同归广泛存在，虽然在某种意义上讲这是正确的，但是在另一种意义上说就不对了。按照惯例来

理解，有了殊途同归性之后，人们将每一个个案分配给产生了结果的某种特殊原因配方(particular causal recipe)。这种做法在定量传统中并不存在。在这里更可以从残差角度来理解个案，即它们在多大程度上靠近那条代表着作为一个整体的因果模型的线。这里没有殊途同归，因为作为一个整体的模型是对总体的解释。

4.6 结论

在两种传承中使用的标准因果模型非常不同，在很多方面难以比较。但是，从两种传承的对话观点来看，最好还是理解其差异而非一争高下。逻辑学和集合论构成了定性因果模型的基础，它们或多或少比作为定量模型基础的概率论和统计学更严格。定性方法强调指出，原因项取决于情境，并且作为整体共同起作用。对于这种研究来说，既然知道其典型因果模型暗示了有少数几条因果路径可达到结果，那么殊途同归就是一个有用的概念。任何结果都有大量的原因存在，这个事实让定量研究看到了因果复杂性。对于这种研究而言，殊途同归并不是一个有用的概念，因为其因果模型同时表明有大量的殊途同归，且仅有一条普遍的因果路径。

参考文献与推荐读物

方程(4.1)表达的各种版本的统计模型都可以在无数统计学教材中找到。该模型也是金、基欧汉和维巴(1994)论证的核心。有关统计学传承中的交互项方面的著述，参见艾利森(Allison, 1977)，布兰博尔、克拉克和戈尔德(Brambor, Clark and Golder, 2006)，弗兰泽塞(Franzese, 2003)以及甘和弗兰泽塞(Kam and Franzese, 2007)。摩根和温希普(Morgan and Winship, 2007)对潜在结果架构的介绍非常好。唐宁(Dunning，即出版)也给出了通俗易懂的引介。有关社会科学中的贝叶斯研究有两项杰出成果，即杰克曼(Jackman, 2009)和吉尔

(Gill, 2007)。在哲学文献中,集合论/逻辑因果模型最早可追溯到麦凯(Mackie, 1980),直到现在(如 Baumgartner, 2009)。在社会科学中,该模型被拉金(Ragin, 1987, 2000)形式化了。有关从统计学角度讨论关于因果复杂性的定性研究,参见阿肯(Achen, 2005),布劳米勒(Braumoeller, 2003, 2004),克拉克、吉利根和戈尔德(Clarke, Gilligan and Golder, 2006),伊莱亚森和斯特赖克(Eliason and Stryker, 2009)以及西赖特(Seawright, 2005)。

Achen, C. 2005. "Two Cheers for Charles Ragin." *Studies in Comparative International Development* 40:27—32.

Allison, P. 1977. "Testing for Interaction in Multiple Regression." *American Journal of Sociology* 83:144—153.

Angrist, J., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics*. Princeton: Princeton University Press.

Bates, D., and D. Watts. 1988. *Nonlinear Regression Analysis and Its Applications*. New York: John Wiley & Sons.

Baumgartner, M. 2009. "Inferring Causal Complexity." *Sociological Methods and Research* 38:71—101.

Brady, H.E. 2010. "Data-set Observations Versus Causal-process Observations: the 2000 U.S.Presidential Election." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E.Brady and D.Collier. Lanham, MD: Rowman & Littlefield.

Brambor, T., W.Clark, and M.Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14:63—82.

Braumoeller, B. 2003. "Causal Complexity and the Study of Politics." *Political Analysis* 11:209—233.

Braumoeller, B. 2004. "Hypothesis Testing and Multiplicative Interaction Terms." *International Organization* 58:807—820.

Cartwright, N. 1989. *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

Clarke, W., M.Gilligan, and M.Golder. 2006. "A Simple Multivariate Test for Asymmetric Hypotheses." *Political Analysis* 14: 311—331.

Downing, B. 1992. *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. Princeton: Princeton University Press.

Dunning, T. Forthcoming. *Natural Experiments in the Social Sciences*. Cambridge: Cambridge University Press.

Eliason, S., and R.Stryker. 2009. "Goodness-of-fit Tests and Descriptive Measures in Fuzzy-set Analysis." *Sociological Methods and Research* 38:102—146.

Esping-Andersen, G. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.

Franzese, Jr., R.J. 2003. "Quantitative Empirical Methods and the Context Conditionality of Class and Modern Comparative Politics." *CP: Newsletter of the Comparative Politics Organized Section of the American Political Science Association* 14:20—24.

George, A., and A.Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Gill, J. 2007. *Bayesian Methods: A Social and Behavioral Sciences Approach*, 2nd edition. London: Chapman and Hall.

Goertz, G. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41: 88—109.

Hicks A., J.Misra, and T.Ng. 1995. "The Programmatic Emergence of the Social Security State." *American Sociological Review* 60:329—350.

Jackman, S. 2009. *Bayesian Analysis for the Social Sciences*. New York: Wiley.

Kam, C.D., and R.J. Franzese Jr. 2007. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: University of Michigan Press.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Lieberman, E. S. 2005. "Nested analysis as a Mixed-method Strategy for Comparative Research." *American Political Science Review* 99:435—452.

Lott, Jr., J.R. 2000. "Gore might Lose a Second Round: Media Suppressed the Bush Vote." *Philadelphia Inquirer*, Tuesday, November 14, p. 23A.

Luebbert, G. 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. New York: Oxford University Press.

Mackie, J. 1980. *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press.

Mahoney, J. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. Cambridge: Cambridge University Press.

Moore, Jr., B. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Morgan, S., and C.Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Ragin, C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of Cal-

ifornia Press.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Schrodt, P. 2010. "Seven deadly Sins of Contemporary Quantitative Political Analysis." Paper presented at the annual meeting of the American Political Science Association.

Seawright, J. 2005. "Qualitative Comparative Analysis vis-à-vis Regression." *Studies in Comparative International Development* 40: 3—26.

Skocpol, T. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.

5 不对称性

对于任何关系来说，……对称性和不对称性之分都是它的（一个）基本维度，尽管实际上大部分人对它完全没有感觉。

——斯坦利·利伯森（Stanley Lieberson）

5.1 导言

定性和定量传承的一个核心区别涉及它们分析对称或不对称关系的倾向。定量学者自然重点关注有对称性质的关系，定性学者则趋于分析有不对称性质的关系。最终，一种关系是否对称，这是一个经验问题。然而，这两种传承的规范促使他们各自考量一种关系而不是另一种。

因果模型和解释可以在各种维度表现出不对称性。在本章中，我们主要（尽管不完全）关注什么可能被称为"静态因果不对称性"（static causal asymmetry）。[①]这个表达意味着对现象发生的解释并不是对其

① 本章也提及因果不对称性（causal assemetry）一词，它涉及这样的观念，即一个变量的变动（比如从有变无）所产生的影响不等于其在其他方向上的变动带来的影响。完全对称的因果效应是这样的，即不管 X 如何改变方向，它对 Y 的效应都相同。我们可以从一种反事实的方面来看待这种因果对称性：当因果关系在反事实的意义上可逆，它就是对称的关系。如果某个给定的 X 返回其初始值，X 对 Y 的影响会消失。相关讨论可参见利伯森（Lieberson，1985：第 4 章）。

未发生的解释的映像。当同一个模型既可以解释发生，也可以解释不发生的时候，就出现了因果对称性。例如，在 logit 模型中，对成功的解释就是对失败的解释的否。实际上，鉴于对 0 和 1 的编码是随意的，因此，对调它们什么都不变。同样，对于全部的 Y 值来说，我们完全可以使用相同的因果模型。相比之下，在因果不对称下，解释出现的模型和解释缺失的模型是不同的。我们不能针对所有的 Y 值用相同的模型。

5.2 对称与不对称模型

为简单起见，现考虑一个 2×2 表。如在第 2 章所见，对于 2×2 表来说，几乎所有的关联测度都是对称的。行或列调换不改变关联的性质。在统计学传承内部，这是一个正向(positive)特征。如果因二分变量的(可能随意地)编码变化而颠覆原有的发现，这是令人沮丧的。事实上，在统计传承中应用的因果性的反事实定义(参见第 6 章)恰恰意味着对称性：人们更感兴趣的是从 0 到 1 的变化，而不是相反。

随着我们从 2×2 表转向连续变量，统计传承中的对称性通常被保留下来(实际研究中的确如此)。例如，在常规最小二乘法(OLS)回归中，人们在估计一个模型时需要假定因果效应是对称的。不管 X 的始点在哪里，也不管 X 的值是增加还是降低，在 X 上给定 1 个单位的变化都被理解为在 Y 上有相同的效应。用线性回归线很容易检视这一点。这条线假定一个完全对称的效应。在 logit 模型中，X 增加 1 个单位对 Y 概率的影响量将等同于下降一个单位的影响。同样，在潜在结果架构中，$Y_t - Y_c$ 项是对称的。

相比之下，集合论模型通常假定不对称关系是围绕必要性和充分性建立的。例如，考虑这样一个关于必要条件的假设。当 $X=0$ 的时候，该假设有一个精确的预言：结果应该不发生。正式地说，必要条件的缺失产生一个点预测：$P(Y=1/X=0)=0$。与之形成鲜明对比的是，当 $X=1$ 时，该必要条件模型却给出一个模糊的说法：$P(Y=1/X$

$=1)>0$。必要条件($X=1$)的出现仅带来$Y=1$出现的“可能性”。所有这些都意味着$Y=1$出现的概率大于0。

当考察Y及其原因的时候，我们看到了相同的基本非对称性。一条必要原因的缺失本身足以解释$Y=0$的个案。相比之下，一条必要原因的出现仅能部分解释$Y=1$的个案。认知心理学中的研究表明，人们倾向于对$Y=0$的个案(而非$Y=1$的个案)给出单因素的必要条件解释。如此看来，当要求被试去解释失败或不发生的时候，他们更可能依靠单变量必要条件的解释。当要求被试去解释成功的时候，他们经常使用更复杂的多变量解释，这种解释可能不援引任何必要条件。

在充分条件下，非对称性则有另类表现：$X=1$会产生一个精确的预言，而$X=0$却做出模糊的声称，即仅声明结果是可能的。[①]这里，一个充分条件的出现就能很好地解释任何$Y=1$个案。相比之下，一个充分条件的缺失则在解释某个个案在Y的值上所言甚少。

随着我们从二分变量转向连续编码的模糊集变量，这些同样的不对称原理仍然适用。例如，在一个必要的原因条件下，较小的X值可以很明确地预测Y的值(即Y将小于或等于X)。当$X=0$这个最小值时，Y必然恰好等于0。随着X值变大，Y的取值域增加。当$X=1$这个最大值时，Y可以取任何值并和假设一致。刻画必要条件的对角线是围绕这种不对称性建立的：Y深受较低的X值的限制，较高的X值却不如此。

5.3 集合论模型的非对称性解释实例

如利伯森(Lieberson, 1985:63)所指出，“大多数经验研究都假定社会关系是不对称的”。对于定量传承来说，这当然为真，它也构成了

① 在社会科学中，很少有人认为单个因素本身就足以带来一个正向结果。相反，人们将多个原因组合在一起才成为某个结果的充分条件。

社会科学中的大量经验研究。然而，在定性传承中，人们也发现很多非对称关系的例子。

表5.1针对非对称关系给出了一个简单易懂的展示（至少对于学术界来说是如此）。这个表包含的是2009年一个一流社会学系的真实招生数据（参见Vaisey，2009）。集合论的解释是，得分高于定量GRE分的中位数（620）就是入学的一个事实上的必要条件（只有一个例外）。这种必要条件的表述可以充分地解释所有得分低于均值的学生为什么不被录取。成绩低于定量GRE分的学生可以根据这个单一变量解释他们为何被拒绝。这当然不是故事的全部，因为（反事实地讲）即便假定这些学生中的大多数已经有很高成绩，也都会被拒绝。大多数人仍然缺乏一套能够被接受的充分条件。在这种意义上，他们的落榜是过度决定了的。不管怎样，较低的GRE数学成绩足以确定他们失败的命运。

通过比较被接纳和被拒绝的学生，我们可以清楚地看到，定量的GRE成绩并不能完全解释被录取。大多数成绩超过均值

表5.1　二分表中的不对称性：某个一流社会学系2009年的招生项目

	GRE分	
	<620	>620
接受	1	34
拒绝	98	209

资料来源：感谢斯蒂芬·韦西（Stephen Vaisey）分享这些数据。也可参见韦西（Vaisey，2009）。

的学生仍然被拒绝了；在将成绩高于均值的考生分为被录取者和被拒绝者的时候，额外的因素在起着作用。①

因此，人们可能询问哪些因素促成了学生被录取。假定我们有一个简单却相当现实的情境：影响录取的因素有考试成绩、GPA、本科毕业学校的水平和推荐信。一个更现实的模型会增加几个附加因素，如

① 当然，高于平均GRE定量分值之间的变异也可能对结果有影响。

论文性文章(writing sample),但是我们的观点可以用这四个因素来表达。利用这些变量,得到关于录取的 logit 模型是:$Y=\beta_0+\beta_1 T+\beta_2 G+\beta_3 U+\beta_4 L+\varepsilon$($T$:考试成绩,$G$:GPA,$U$:本科毕业院校水平,$L$:推荐信)。一个可选择的集合论模型是:$Y=T*G*(U+L)$。这个模型给出了如下对失败的解释:$\neg Y=\neg T+\neg G+(\neg U*\neg L)$。因此,导致失败的因素可能是低分或低水平学校或本科学校低质量和印象不深刻的推荐信的组合。

就 logit 模型来说,应注意其中如何有一种关于成功和失败的单因素解释:原因项的加权总和或者足够高,以至于跨越某个成功的门槛,或者太低而导致失败。[①]整个模型即涉及将 0 个案和 1 个案进行比较。我们不能设想针对 $Y=1$ 和 $Y=0$ 分别写出 logit 模型。[②]相比之下,在集合论研究中,关于成功和失败的解释(尽管相关并使用相同的因素)却有不同的形式。有两条通向成功的路径(即 $T*G*U$ 或 $T*G*L$),而通向失败的路径却有三条(即$\neg T$, $\neg G$, $(\neg U*\neg L)$)。为求成功,至少需要考虑三个因素,而一个因素通常就足以导致失败。这看上去与实际录取程序一致:为了迅速淘汰掉大多数申请者,委员会成员要寻找致命的缺陷,然而他们要考虑一系列因素并仔细审阅才能确定成功者的申请者。

这个简单的例子提请我们注意基于逻辑学的模型有一个重要性质:

> 解释的不对称性。对失败的解释通常不同于对成功的解释。

一个失败结果的原因未必等价于成功结果的原因的缺失或否定。事实上,在使用形式布尔方法的定性研究——如定性比较分析(QCA)

① 也可以建立二分变量的 logit 模型,其中只有少数几条路径可带来成功,导致失败的路径的数目不同于导致成功的路径数[感谢撒德·邓宁(Thad Dunning)指出这一点]。显然,在一个 logit 模型中,可能带来成功的路径数将取决于变量和参数估计的性质。然而,当利用不带极端值的连续变量的时候(通常还真这样做),将存在多条通向成功的路径。

② 的确存在前设不对称性的统计模型。例如,马尔科夫转换矩阵(Markov transition matrices)不假定从状态 i 转到状态 j 的概率等于反向的概率。

中，研究者对于成功和失败的个案通常得出相当不同的最终模型(Ragin，2000)。下面的三例说明了这一点：

1. 威克姆-克劳利(Wickham-Crowley，1992)探索了拉丁美洲成功的革命游击运动的原因。他的布尔分析表明存在如下五个条件，其中每一个单独都是社会革命成功的必要条件，合起来则是其充分条件：(A)游击战尝试；(B)农民的支持；(C)强大的游击队；(D)世袭的禁卫军制度；(E)失去美国的支持。因此，他给出的成功因果模型是：

$$\text{成功的社会革命} = ABCDE$$

相比之下，威克姆-克劳利对失败的革命游击运动的解释是非常不同的：

$$\text{社会革命的失败} = ABd + bce + bcD$$

这种非对称性是显而易见的：通向成功的路径只有一条，它含有五个要素，而导致失败的路径却有三条，且每一条都包含三个因素。值得注意的是，某些个体因素的因果效应大大依赖于情境(即与它们相结合的其他因素)。例如，原因项(causal factor)B(农民支持游击队)对社会革命而言是必要的，而当游击队进攻的是一个非世袭禁卫军政权时，它却有助于防止社会革命的发生。我们也可以用下列方式思考这些方程的非对称性：Abd 的确足以确保不发生社会革命，然而反之却不成立。也就是说，否定这些因素也不会促成一场社会革命。

2. 斯托克(Stokke，2007)考察了哪些因素可以导致当事者遵循国际渔业体制。他给出五个原因项：(A)建议——来自体制内科学咨询机构的明确建议；(C)承诺——制度受众的行为明确违反了保护性措施；(S)预期考量(shadow of the future)——制度受众在现行体制下追求新交易的感知性需求；(I)不便之处——制度受众的行为不便于改变；(R)反响——受众冒着因不服从而受诽谤的风险。斯托克给出的最终成功模型如下：

$$成功 = Ai + ARS$$

因此,通往成功的路径有两条,明确的科学建议(A)一般是成功遵循国际渔业体制的必要条件。人们可以运行斯托克的数据并生成一个失败(即非成功)模型:

$$失败 = sI + Ir$$

此外,失败方程并不简单地是对成功方程的否定;这些方程不是对称的。例如,在成功方程中,“建议”因素是一个必要条件,但它始终没有出现在失败方程中。

3. 马奥尼(Mahoney, 2010)认为,西班牙殖民主义的变化可以解释拉丁美洲诸国之间的经济长期发展水平的差异。在最聚合的层次上,他的解释强调了三种原因项:(M)重商主义的殖民中心——国家深受西班牙哈布斯堡王朝(Habsburg Spain)(1500—1700)的统治;(L)自由殖民中心——国家深受西班牙波旁王朝(1700—1821)的统治;(W)福利——国家在后殖民时期经历了代价高昂的战争。他的综合发现可归纳如下:

$$高度经济发展 = mL + mlw$$

$$中度经济发展 = ML$$

$$低度经济发展 = Ml + mlW$$

显然,这些解释之间不存在完美的对称。通往高度经济发展的路径有两条:(1)一个国家在重商主义哈布斯堡王朝的殖民主义阶段处于边缘地域,而在自由主义波旁王朝时期处于核心地区;(2)一个国家可能在这两个时期都处于边缘地域,随后在后殖民主义时期避免了代价高昂的战争。通向中度经济发展的路径只有一条,即国家在重商和自由主义阶段都处于殖民的中心。对于这条路径来说,代价高昂的后殖民战争是否发生却无关紧要。最后,达至低水平经济发展的路径同样有两条:(1)国家在重商主义哈布斯堡王朝时期处于核心地区,在自由主义波旁王朝阶段处于边缘地域;(2)国家在这两个时期都处于边缘地域,之后在后殖民主义时期经历了高代价的战争。

最后，为了说明寻求非对称关系有什么潜在价值，让我们考虑霍华德和勒斯勒尔(Howard and Roessler, 2006)的获奖文章，“竞争性威权政权下的自由选举结果”。[①]因变量是政权的自由化程度，未必是民主化，但是朝此方向运动。我们和他们都关注的关键自变量是“反对者联盟”。这意味着，反对政府执政的团体和政党在即将到来的选举中已经形成了一个联盟(分析单位是选举)。表 5.2 再现了他们的统计结果。

表 5.2 的统计发现清楚地显示了反对者联盟变量的强烈影响。这可以说是表中最强的一对关系，也基本上是霍华德和勒斯勒尔最中意的一个变量。同样真实的是，该变量同时解释了成功与失败：缺席与失败相关，在场与成功相伴。就本章的内容来说，我们可以追问，反对者联盟变量在解释结果的时候是否不对称。换言之，反对者联盟存在的影响与其不存在的影响是否不同？

表 5.2　一个对称的统计模型：对自由选举结果的解释

自变量	参数估计
反对者联盟	7.8** (3.0)
反对者动员	0.91* (0.4)
推翻当政者	3.2* (1.5)
经济增长	0.3(0.2)
对外直接投资	−0.1(0.3)
对外援助	0.01(0.02)
议会制	−3.3(2.2)
政权的开放度	1.0(1.0)
先前的自由式变革	−1.4(1.7)
常　量	−1.3(5.2)
总数 *N*	50

注：括号内为标准误。

资料来源：Howard and Roessler, 2006:375。

① 感谢卡斯滕·施耐德(Carsten Schneider)帮我找到这个例子并允许我们使用它。参见施耐德(Schneider, 2007)对这些数据的充分再分析。

通常情况下，所有读者都知道统计结果表（即表 5.2）中蕴涵的内容。为了探讨其中是否存在因果不对称性，一种做法是做一个全面的定性比较分析（QCA）（参见 Schneider, 2007）。然而，为求简单，现考察一个 2×2 表（参见表 5.3）。正如我们在第 2 章中所见，通过集合论来考察数据所导致的解释可能不同于依据统计分析给出的解释。

一方面，很容易看出为什么人们可能认为这个表是有关因果对称性的一个很好例证：73%（8/11）的反对者联盟个案引起了自由化，82%（32/39）的非反对者联盟个案导向了非自由化。因此，反对者联盟是否存在似乎能以一种相当对称的方式来预测是否自由化。

另一方面，如果透过定性视角来看，我们的注意力就转向那些几乎空的格。这种看待数据的方式唤起我们关注潜在的不对称性。纵观各行，我们发现反对者联盟的缺失近似成为自由化缺失的一个必要条件，即 91%（32/35）的非自由化个案也缺少反对者联盟。这表明，反对者的缺失在解释失败时至关重要。①

表 5.3　统计分析与不对称因果：对自由选举结果的解释

		反对者联盟 0	反对者联盟 1
LEO	1	7	8
LEO	0	32	3

注：LEO 表示自由选举的结果（liberalizing electoral outcome）。
资料来源：Schneider, 2007。

在实现自由化的诸多个案中，近 50%（7/15）设法在没有反对者联盟的情况下获取成功。因此，反对者联盟的存在对成功而言不再是必要条件，也不再接近于是充分条件。看起来，存在着多条实现自由化的途径，其中都不包括反对者联盟的存在，但是，通往失败之途却几乎总是包含着反对者联盟的缺失。因此，我们对这些数据的稍加分析遵循着与社会学招生数据同样的形式。我们有一个变量在解释失败时比解

① 一个更深入的分析将探讨它在何种程度上是一个不足道的必要条件或具有低覆盖度（Goertz, 2006; Ragin, 2008）。

释成功更重要。

显然，这个例子仅仅是说明性的。我们还没有考察控制变量和其他可能导致一个不同解释的因素。关键在于，被认为是数据的最重要的那些特征可能变化极大，这取决于人们调换到对称模式还是不对称模式。虽然数据本身不变，但得出的推断却取决于人们的视角。

5.4 结论

居于两种传承核心的基本因果模型在对称性问题上存在分歧。集合论模型处理的是不对称关系。他们假定不同的 X 值会对 Y 产生不同的效应。同样，他们通常针对不同的 Y 值（如成功与失败）给出不同的解释。相比之下，统计模型（恰如在实际研究中的应用那样）通常是对称的。线性回归模型假定不管 X 的起点值怎样，X 的一个给定单位的变化都将对 Y 产生同样的影响。在 logit 模型中，不管因变量的值由 0 变成 1 还是反之，都会产生同样的参数估计值，除非正负号改变。由此，解释不发生是解释发生的映像。

我们并不主张一种路数优于另一种。最终而言，对称性与不对称性的存在取决于数据和有待检验的因果模型。然而，我们的确相信，这种差异是为何难以比较定性和定量因果模型的关键原因。很难将标准统计模型的基本对称性转译为集合论模型的基本不对称性。类似地，也难以用标准的、对称的统计学工具来把握集合论模型的不对称性。

参考文献与推荐读物

方法论文献通常不考虑统计模型的静态因果对称性，尽管偶尔有深入考察（例如，Abbott，2001：第 1 章）。对定性比较分析的标准的讨论强调，作为方法论的一个常规部分，人们需要分别估计成功模型（$Y=1$）和失败模型（$Y=0$）（例如，Ragin，2000，2008；Schneider and Wagernann，forthcoming）。关于对统计模型的动态因果对称性的讨论和批

判,参见利伯森(Lieberson, 1985:第 4 章)。金、基欧汉和维巴(King, Keohane and Verba, 1994, 89—91)简要地回应了利伯森。

Abbott, A. 2001. *Time Matters: On Theory and Method*. Chicago: University of Chicago Press.

Goertz, G. 2006. "Assessing the Trivialness, Relevance and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41: 88—99.

Howard, M., and P.Roessler. 2006. "Liberalizing Electoral Outcomes in Competitive Authoritarian Regimes." *American Journal of Political Science* 50:365—381.

King, G., R. Keohane, and S. Verba. 1994.*Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Lieberson, S. 1985. *Making It Count: The Improvement of Social Research and Theory*. Berkeley: University of California Press.

Mahoney, J. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. Cambridge: Cambridge University Press.

Ragin, C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Schneider, C. 2007. "Opposition in Transition. does Unity Bring Dictators Down? A QCA Re-analysis of Howard & Roessler's Study of Liberalizing Electoral Outcomes." Manuscript, Central European University.

Schneider, C., and C. Wagemann. Forthcoming. *Set-theoretic Methods in the Social Sciences*. New York: Cambridge University Press.

Stokke, O. 2007. "Qualitative Comparative Analysis, Shaming, and International Regime Effectiveness." *Journal of Business Research* 60:501—511.

Vaisey, S. 2009. QCA 3.0: "The 'Ragin Revolution' Continues." *Contemporary Sociology* 38:308—312.

Wickham-Crowley, T. 1992. *Guerrillas and Revolution in Latin America: A Comparative Study of Insurgents and Regimes Since 1956*. Princeton: Princeton University Press.

6 休谟对原因的两种界定

有之，显“效”；无之，“效”散；无以名“之”；唯曰“要因”。

——伽利略

大多数原因都可以更准确地被认为是“充分非必要条件的必要非充分部分”(INUS conditions)。

——威廉·沙迪什(William R. Shadish)

托马斯·库克(Thomas D. Cook)

唐纳德·坎贝尔(Donald T. Campbell)

6.1 导言

定量和定性传承中的因果关系有不同的研究路数。来自大卫·休谟的一段著名引文为我们提供了讨论这些路数的一种有效的方式：

> 我们可以将原因定义为一个物象随后的另一个物象，在这里我们可以说，凡是和后者相似的一切物象都必然伴随着和前者相似的物象(定义1)。或者，换言之，如果第一个物象不存在，第二个物象将永远不会存在(定义2)。(大卫·休谟：《人类理解和道德原则研究》)

正如许多哲学家所建议的那样，休谟的短语“换言之”是有问题的。该短语使得定义 1 与定义 2 好像是等价的，而事实上它们代表完全不同的路数。刘易斯写道，“休谟的‘换言之’……不仅仅是重述了他的第一个定义。它们提出了完全不同的东西：对原因的一种反事实分析”（Lewis，1986a：160）。

遵循刘易斯的分析，我们可以称休谟的定义 2 为“反事实的定义”。相比之下，为了突出休谟的观念，即原因后面总是紧随着结果，我们应该称定义 1 为“恒随定义”（constant conjunction definition）。[①]在本章中，我们思考在定量和定性研究传统中，这两种定义是如何与对因果关系的理解相关的。

需要强调的是，我们在运用休谟的理念的时候，仅将它作为一个手段，用它来探讨“原因”概念可通过哪些不同的方式用于定量和定性传承中。我们关注的是，在哲学中对原因的潜在理解能有助于推动两种范式的发展。还值得注意的是，我们的理解不应归因于休谟本人。休谟关于因果性的观点已经成为哲学家之间大量争论的来源，我们也不声称去解决这些争论。

6.2 定量传统

在潜在结果架构（potential outcome framework）兴起之前（参见 Morgan and Winship，2007）就有对因果关系的统计学讨论，这种讨论关注的是在一种概率论架构下的休谟的恒随定义（定义 1）。例如，苏佩斯在其早期的著名分析中写道：“粗略地说，我对休谟的分析提出如下修正，即如果第一个事件出现之后，伴随而来的第二个事件出现的概

① 这种因果观支撑了 20 世纪中期社会科学中形成的覆盖律模型（the covering law model）。例如：“[一个覆盖的、科学的（Hempel，1965）]定律有如下形式‘如果存在条件 C_1，C_2，…，C_n，就总有 E’”（Elster，1999：5）。

率很高，则第一个事件是第二个事件的原因。”(Suppes, 1970:10)[①]在这种概率论的路数中，似乎自然地基于关联性来理解恒随定义：$X=1$ 与 $Y=1$ 有关。

针对休谟的反事实定义(定义 2)，我们也可以提出一种统计学解释。然而，这需要某些准备工作，因为休谟的反事实定义暗指单个个案。定义 1 陈述的是“所有对象(复数)后面都紧随……”，与之不同，定义 2 阐述的是“如果第一个物像(单数)不存在……”因此，在一个恒随方式中来理解定义 2，就要求将休谟的观念扩展到多个个案中。

定量传统从跨个案的恒随性(constant conjunction)的角度解释定义 1 和定义 2，由此实现了这一扩展。相关系数为 1.00 意味着 $X=1$，$Y=1$ 和 $X=0$，$Y=0$ 之间具有一种恒随性。因此，可以将定义 1 和定义 2 融合成一个统计解释。定义 1 认为，当原因存在时，结果将(在概率意义上)存在。定义 2 认为，当原因不存在时，结果将(在概率意义上)不存在。由于仅仅考察 $X=1$ 而缺少 $X=0$(反之亦然)的个案是没有统计意义的，这两种定义便统合成了一个。二者都不能单独存在并具备统计意义。但当它们融合成一体，便提供了一个对因果性的连贯对称性的理解，其中的重点是分析在自变量取不同值时结局如何。

目前，在政治科学和社会学的方法论者中间，探讨因果关系时的一个主要观点就是潜在结果架构。也许它在统计学界的最重要创新是强调因果关系的反事实基础。例如，摩根和温希普(2007)的杰出综述名为《反事实和因果推断》(*Counterfactuals and Causal Inference*)。早期的统计学和概率论解释都被认为忽视或低估了因果关系的这一重要方面。

Neyman-Rubin-Holland 模型处于该路数的核心，该模型首先始于个案，然后建构一个成熟的因果关系统计模型。该模型运用基本的实验设计，对个体 i 进行实验处理。反事实就是：如果 i 接受了控制后会

① 显然，苏佩斯的解释比较复杂。尤其是当处理观测数据时，(与处理实验数据相反)人们必须关注虚假关系等。

发生什么。由于个体不能同时既接受处理又受到控制，两种可能之一必然总是一个反事实。这一实情引出了一个基本问题：

> 因果推断基本问题。不可能在同一单位中观测到 $Y_t(i)$ 和 $Y_c(i)$ 的值（t = 处理，c = 控制），所以，不可能观测到 t 对 i 的影响。(Holland, 1986:947)

由于用统计量来估计或评估因果效应需要有相对大量的实际数据，统计学家最多能做到的就是估计平均因果效应（average causal effect），或者使用比较流行的术语即样本中的平均处理效应（Average Treatment Effect, ATE）。

> 重要的是，统计的解用 t 在整个单位总体上可以估计的平均因果效应代替了 t 在具体单位上不可能观测到的因果效应。(Holland, 1986:947)

将获得的平均处理效应（ATE）作为因果关系反事实理论的基础，这便完成了定量传承中我们所谓的"因果推断循环"（causal inference circle）（类似于解释学循环）：

1. 人们始于休谟的定义 2，它强调针对被试 i 的反事实。

2. 运用代数和统计量来解释定义：反事实是处理和控制之差，即 $Y_t(i)-Y_c(i)$（Holland, 1986:947）。

3. 人们用恒随形式的定义 1 来分别进行处理和控制，即分别针对 $X=1$ 和 $X=0$ 进行处理和控制。

4. 在所有的个案中，计算出处理和控制之间的平均差，即平均处理效应（ATE）。

5. 平均处理效应便为被试 i 提供了个体个案的反事实。

在这一循环中，休谟的恒随定义 1 是举足轻重的，尽管其出发点是他的反事实定义 2。反事实的出发点增加了解决问题的不可能性。其结果是，学者必须迅速转向定义 1，并将恒随观念既用在处理，也用于

控制，从而取得进展。然而，其结果是这种架构在估计因果效应的实践中遵循了先前的统计路数，即将反事实的定义 2 还原到恒随定义 1。

6.3　定性传统

从定性传统的视角看待事物，就可以从逻辑学方面来理解休谟的定义。哲学家和定性方法论者关注恒随定义的逻辑形式："若 $X=1$，则 $Y=1$。"将"如果……那么"分句作为一条逻辑陈述来理解，定义 1 就将"原因"看成是 X 与 Y 之间的一种充分性的关系。这一充分性解释关注的是 $X=1$ 的情况（即那些存在着原因的个案）。研究者从存在的原因出发，然后考量它是否有相应的效应。在此意义上，对定义 1 的定性解释与定量解释类似。

然而，在这一点上，两种传统分道扬镳。定性解释不会做出进一步推断，即认为休谟的恒随定义暗含着 $X=0$ 与 $Y=0$ 的个案之间有相关性。相反，这种观点认为，如果原因不存在，其结果可能存在或不存在。定性解释将定义 1 作为一个关于充分性的声称，这一充分性可以被独立研究。因此，虽然统计学解释将定义 1 和定义 2 融合起来，但与之不同，对定义 1 的充分性解释完全自行成立且可能是有效的，即独立于当使用定义 2 时得到的结论。

两种传统之间在定义 1 上的另一个重要差异涉及这一事实，即效应很少总是由个体的诸多原因带来的（至少在社会科学中是如此）。正如我们在上文看到的，定量方法早已用概率论假定解决了这一问题。虽然这些假定有时也被纳入定性研究中，但是另一标准的解是将因果充分性与"多重的、联合的"因果关系连接起来（Ragin，1987）。正如在第 4 章"因果模型"中所讨论的，定性研究者将因果关系视作组合性的或"联合的"，其意义是说，几种不同的原因必须组合起来（如 $X_1 * X_2 * X_3$）才能产生一个结果。至于个体原因（如 X_1）本身不足以带来结果。

同时，定性研究者将因果充分性视作具有潜在的“多重性”，即通常存在着多因素的“不同组合”，每一个组合对结果来说都是充分的。这便是一般的殊途同归原则：没有任何单个原因序列（package of causes）可以产生所有 $Y=1$ 的结果。不同的原因序列会导致同样的结果。在定义 1 的情境下，这一理念是至关重要的；它暗指的含义是，在给定的一个充分性序列上的零值并不意味着 $Y=0$，因为其他因果序列可能会产生 $Y=1$ 的结果。

这种基于逻辑学的路数生成了自己的推理链条，它始于休谟的定义 1，达至个案：

1. 始于休谟的定义 1，它强调恒随性。

2. 对该定义的解释意味着 $X=1$ 对于 $Y=1$ 来说是充分的。

3. 将 X 视为是由一序列原因项组成的，如 $X_1 * X_2 * X_3$。

4. 推而广之，认为所有原因系列 $X_1 * X_2 * X_3$ 之后都“跟随着” $Y=1$。

5. 如果个案 i 包含 $X_1 * X_2 * X_3$，那么这一序列就被理解为是 $Y=1$ 的原因。①

因此，我们结束于关于个案 i 的因果陈述，就像我们之前针对统计因果推断循环所做的一样。尽管二者有这种相似性，但是在推理过程中使用的步骤是截然不同的，而最终对于 i 的因果断言的本质也明显不同。统计路数同时使用定义 1 和定义 2，而这种版本的定性路数只使用定义 1。

也可以从逻辑的角度来看待休谟的反事实定义 2。在哲学界，因果关系的反事实方面一直受关注。可以说，哲学界近几十年来对因果关系的最有影响力的解释来自戴维·刘易斯。他的专著《反事实》（*Counterfactuals*）最先于 1973 年出版，早于统计学中对反事实的重新发现。与休谟的定义 2 相一致的是，刘易斯根据个体个案提出了他的反事实定义：

---

① 最后关于 i 的推理要假定第四步中的推广是有效的。我们先前假定存在着显著的平均处理效应，这两种假定类似。

> 我的分析意味着要应用于特定个案中的因果关系。它不是对因果一般性的分析。(Lewis, 1986a:161—162)
>
> 如果特定事件 c 不发生,事件 e 就不会发生;当且仅当此时,才称事件 e 因果地依赖于事件 c。(Lewis, 1986a:242)

统计学以外的其他著作也强调了个体个案中的因果关系,包括马克斯·韦伯对反事实的著名分析(Weber, 1949),哈特和奥诺雷(Hart and Honoré, 1985)对法律中因果关系的分析。当关注点是个体事件时,反事实的解释便是自然而然的选择(包括在潜在结果架构中亦如此)。

大多数有关反事实的哲学文献仍然维持在单个个案层次,但社会科学的核心目标在于普遍解释,因此,如何用因果规律性(causal regularities)来改写定义 2,这是社会科学家特别感兴趣的课题。在定性传统中,关键的进展是用一个必要条件来解释反事实。因此,人们可以利用逻辑学,根据下列方式重述定义 2:如果 $\neg X_i$,则 $\neg Y_i$。这似乎是完全自然的,因为休谟说:"如果第一个物象未曾存在。"如果关注点又回到众多个案和一般模式,在此意义上,有关必要原因的假设将我们带回到休谟的恒随定义 1。

定性研究者将反事实进行推广,推广的过程启示了另一个始于定义 2 的因果推论循环:

1. 始于休谟的定义 2,它强调的是反事实。

2. 从逻辑角度来理解这一定义:如果 X 没有发生过,Y 也不会发生,即如果 $\neg X_i$,那么 $\neg Y_i$。

3. 将个别个案的反事实推广到所有个案。即对于所有的 i 来说,如果 $\neg X_i$,那么 $\neg Y_i$。[①]

4. 运用定义 1,将这种反事实转换为关于一个必要原因的普遍表述,即 X 对于 Y 来说是必要的。

5. 如果 X 在情况 i 中存在,那么 X 是 Y 的一个原因。

---

① 当然,"如果 $\neg X_i$,那么 $\neg Y_i$"与"如果 $Y = 1$,那么 $X = 1$"是等价的。但是,这一表述逆转了反事实版本的因果方向。虽然如此,"如果 $Y = 1$,那么 $X = 1$"的表述对于经验检验而言仍然是重要的,因为人们往往用这种表述来选择个案(参见第 14 章)。

在这一循环中，关键的进展是将个别个案的反事实转换为关于一个必要原因的规律性陈述。实际上，分析者在循环中始终与定义 2 相伴，引入定义 1 的目的是便于产生跨个案的推广。要想保留定义 2，需要假定该定义可以被直接扩展到许多个案中，由此可实现推广。

6.4 结论

休谟的著名引文包含对因果关系的两种定义。定义 1 表明一个介于原因和效果之间的恒随性，从而使得一种效应总是紧随原因之后。该定义假定存在着多个个案，并与定量视角下的因果关系具有亲和性。定义 2 表明因果关系的一个反事实观点，其中原因的缺失将导致结果的缺席。这一定义建立在单个个案之上，并与定性视角下的因果关系具有重要联系。

虽然定量学者似乎自然更倾向于定义 1，但是近些年来，随着人们将注意力转向反事实，定义 2 也在统计方法论者中间成为定义因果关系的起点。然而，定量路数很快就放弃了适用于个案的反事实因果观，并认为不可能估计这种因果关系。在统计学传承中，真的不存在两种不同的定义，因为每个定义单独存在都没有意义。统计路数将两种定义融为一种，从不可能估计定义 2 转变为可能估计平均处理效应（ATE）。

在定性传统中，这两种定义仍然独立。定义 1 被理解为代表着关于因果充分性的声称，而定义 2 被理解为代表了关于必要条件的声称。其结果是，不同群体的学者倾向于其中一种定义而不是另一个。一些学者运用诸如定性比较分析这样的方法来检验因果充分性，这些学者自然地倾向于定义 1 和充分性研究（Ragin，1987）。相比之下，其他定性学者则探讨关于必要原因的假设，他们可能更自然地信奉反事实定义 2（Goertz and Starr，2003）。然而，这两种关于因果关系的定义容易共存，因为它们根植于相同的逻辑学与集合论传统。

是否存在对休谟的两种定义的正确理解呢？我们虽然不是哲学史

家，但认为人们所持的对休谟观点的最有用的理解将受其方法论背景和路数的强烈影响。我们认为，与两种传承之争相一致的是，每一种理解在其嵌入的总的传统中都有其道理。

参考文献与推荐读物

休谟对原因的探讨主要贯穿在关于因果性的哲学文献之中。很多具有影响力的研究都以此为出发点，如苏佩斯(Suppes，1970)的概率因果论、刘易斯(Lewis，1973)对反事实的因果定义、麦凯(Mackie，1980)对"充分非必要条件的必要非充分部分"(INUS)原因的探讨、萨蒙(Salmon，1998)提出的因果关系本体观或机械观，还有近期鲍姆加特纳(Baumgartner，2008，2009)的研究。有关逻辑学(如 Copi，1982)和法律中因果性的著名研究(Hart and Honoré，1985)同样以休谟为中心。对休谟观点的经典处理是比彻姆和罗森堡(Beauchamp and Rosenberg，1981)。在社会科学中，方法论者较少关注休谟对因果关系的论述，虽然这一工作在霍兰(Holland，1986)，格尔茨和斯塔尔(Goertz and Starr，2003)，尤其在 Brady(2008)中有所讨论。在 Holland 之后有 Neyman(1923/1990)和 Rubin(1974；1978)开展的一些核心研究工作。这些工作组成了 Neyman-Holland-Rubin 做出的因果性定义(ATE)，该定义已经在本章中有过讨论。也可参见金、基欧汉和维巴(King，Keohane and Verba，1994)，摩根和温希普(Morgan and Winship，2007)。

Baumgartner，M. 2008. "Regularity Theories Reassessed." *Philosophia* 36：327—354.

Baumgartner，M. 2009. "Inferring causal complexity." *Sociological Methods & Research* 38：71—101.

Beauchamp，T.L.，and A. Rosenberg. 1981. *Hume and the Problem of Causation*. Oxford：Oxford University Press.

Brady，H. E. 2008. "Causation and Explanation in Social Science." In *The Oxford Handbook of Political Methodology*，edited by J. M. Box-Steffensmeier，H. E. Brady，and D. Collier. Oxford：

Oxford University Press.

Copi, I.M. 1982. *Introduction to Logic*, 6th edition. New York: Macmillan.

Elster, J. 1999. *Strong Feelings: Emotion, Addiction, and Human Behavior*. Cambridge: MIT Press.

Goertz, G. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41:88—109.

Goertz, G., and H. Starr, eds. 2003. *Necessary Conditions: Theory, Methodology, and Applications*. New York: Rowman & Littlefield.

Hart, H.L.A., and T.Honoré. 1985. *Causation in the Law*, 2nd edition. Oxford: Oxford University Press.

Hempel, C. 1965. *Aspects of Scientific Explanation*. New York: Free Press.

Holland, P. 1986. "Statistics and Causal Inference(with Discussion)." *Journal of the American Statistical Association* 81:945—960.

Hume, D. 1975(1777). *Enquiries Concerning Human Understanding, and Concerning the Principles of Morals*. Oxford: Oxford University Press.

King, G., R.Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Lewis, D. 1973. *Counterfactuals*. Cambridge: Harvard University Press.

Lewis, D. 1986a. "Causation. Postscripts to 'Causation.'" *Philosophical Papers, vol.2*. Oxford: Oxford University Press.

Lewis, D. 1986b. "Causal Explanation." *Philosophical Papers, vol.2*. Oxford: Oxford University Press.

Mackie, J. 1980. *The Cement of the Universe: A Study of Cau-*

sation. Oxford: Oxford University Press.

Morgan, S., and C. Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Neyman, J. 1923/1990. "On the application of probability theory to agricultural experiments: essay on principles." *Statistical Science* 5:465—472.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Ragin, C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Rubin, D.B. 1974. "Estimating Causal Effects of treatments in randomized and nonrandomized studies." *Journal of Educational Psychology* 66:688—701.

Rubin, D.B. 1978. "Bayesian Inference for Causal Effects: the Role of Randomization." *The Annals of Statistics* 6:34—58.

Rubin, D. B. 1990. "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5:472—480.

Salmon, W.C. 1998. *Causality and Explanation*. Oxford: Oxford University Press.

Schneider, C., and C. Wagemann. Forthcoming. *Set-theoretic Methods in the Social Sciences*. New York: Cambridge University Press.

Suppes, P. 1970. *A Probabilistic Theory of Causality*. Amsterdam: North Holland.

Weber, M. 1949. *Max Weber on the Methodology of the Social Sciences*. New York: Free Press.

第二部分　个案内分析

7 个案内对跨个案的因果分析

个案内比较对于小样本分析的可行性来说是关键。

——戴维·科利尔(David Collier)

7.1 导言

定性和定量传统有一个最根本的区别,它涉及个案内分析和跨个案分析在因果推断方面哪一个更重要。在定性研究中,人们总是重点关注在个案内部发生的特殊事件和过程。关于假设检验的一些重要的定性方法论,如过程追踪和反事实分析,基本上都是个案内分析法。为了使用这些方法,定性分析者必须在其个体个案内部获得关键的观测信息。

定性研究也经常包括一个跨个案的成分。无论对于社会科学中比较常见的小样本研究来说,还是不太常见的中等样本的定性研究而言,这都是成立的(参见附录)。尽管可以通过增加样本的规模来获得一些影响力,但是如果个案的总规模还是很小的话,那么因果推断的主要基础就必须源于个案内分析。小样本比较通常不保障对假设实施有效力的跨个案检验。[1]只有当定性研究的样本量增加,超过少量个案的时

---

① 某些检验的设计目的是为了消除那些设定有潜在的必要或充分原因的假设,小样本比较的确可允许诸如此类的检验。

候，才可能用跨个案方法进行有力的假设检验。

相比之下，定量研究在传统上就包含着专门的跨个案比较。由于它们处理的是大样本数据，定量学者通常根本不了解个案。例如，在调查研究中，学者除了知道受访者对特定问题的回答以外，对做答的个人几乎一无所知。当学者利用大样本的国家数据时，除了所测量的变量之外，对许多国家都所知甚少。由于人力和资源的限制，期待大样本的研究者对他们的大多数个案都有深入的理解是不现实的。

然而，由于定量学者可以就诸多特定的变量来系统地测量和比较个案，所以他们可以并且的确通过跨个案的检验引出一些推断。原则上讲，他们分析的个案数量大，可以有力地检验一些发现，这些发现可不单纯是随机的产物或者干扰变量的结果。定量研究者倾向于把他们的大样本跨个案研究和深度的个案研究结合起来，就此而言，他们走向了一种不同的研究设计——混合方法研究设计。当追求混合方法研究的时候，定量分析者便在单独的一项研究中将跨个案研究和个案内分析结合起来了。

不管怎样，总体而言，小规模定性推断主要依赖于个案内分析，而大规模定量推断主要依赖于跨个案分析。由此，我们可以阐述这两种范式之间的差异，这也是本章明确地关注的任务：

在小规模的定性研究中，因果推断的主要手段源自个案内分析，跨个案的方法论有时候发挥一定的支持作用。

在大规模的统计研究中，因果推断的主要手段源自跨个案分析，个案内方法论有时候发挥一定的支持作用。

个案内与跨个案分析的差异会通过本书讨论的许多对比呈现出来。还有一些重要成果——如乔治和贝内特（George and Bennett，2005），科利尔、布雷迪和西赖特（Collier，Brady and Seawright，2010）——探讨了定性和多元方法研究，提炼了定性和定量研究的对比。个案内与跨个案分析的差异也居于这种对比的核心。

7.2 个案研究/小样本研究中的个案内分析对比实验/统计研究

就个案研究相对于实验来说,简单地对比个案内分析和跨个案分析的典型角色(或非角色)是有用的。首先,考虑一个解释性的个案研究,其中要解释为什么一个特殊的个案会引发一个特别的结果。根据定义,个案研究主要聚焦于一个单独的个案,因此,跨个案分析不是其核心的推断模式。在个案研究中,也可能或明或暗地与其他个案进行对比和比较。此外,当利用过程追踪和反事实分析来检验假设的时候,研究者会利用来自其他个案的公认的概括和发现。然而,根据定义,个案研究主要是对单个个案进行持续的分析。

在努力阐释一项好的解释时,个案研究者会不可避免地采取一种历时性的、过程性的个案分析。其中要考虑到在不同时间点上的许多不同的观测结果。当一些关键事件使个案走向某些结果而不是其他结果时,分析者就通常会确定这些历史关节点。他或她可能停下来,去考虑在这些节点上的微小变化如何可能促使该个案朝着不同的途径发展。也可能考虑随着时间的推移,有历史意义的事件在哪些方面一个个地联系在一起最终产生一系列原因,进而带来所感兴趣的结果。这些全部的解释将可能是丰富的,能够了解有关特殊事件、关节点和意外事件的细节。

现在考虑这样一项实验研究,它试图估计某种感兴趣的处理所带来的效应。在尝试阐述一个有效答案的过程中,通常要进行完全跨个案的分析。研究者会努力分离出所关注的处理带来的效应。针对大样本的个体(或者其他单位),要对实验组和控制组进行随机分配,这在理想状态下有助于抵消历史的和所有其他混淆原因的效应。设计的实验恰恰可以告诉我们,在去掉所有其他因素(包含环境和历史)的效应之后,有哪些纯的因果效应。在这个意义上,实验基本上就是跨个案的设计。

在个案研究比对随机研究中使用的个案内分析和跨个案分析之间存在着差异。在当代政治科学和社会学领域,通常践行的是小样本的

定性研究和大样本的统计研究。当我们进入这些研究之时，上述差异便被大量地再生产出来。小样本定性研究者仍然最关心追踪个案内的过程分析，以便解释特殊的结果。虽然许多定性研究使用了跨个案分析并且经常利用简单配对法（如穆勒的求异法），但是其叙事仍然围绕着个案内过程来实施。实际上，除非研究的个案量大到一定程度，否则，不现实的想法是认为这些小样本比较方法本身能够为大多数因果推断提供强有力的基础。当我们从一个个案转向三、四个个案时，要是没有任何个案内分析，为了检验解释所用的手段就没有说服力。个案内分析必须在假设检验中担当重任。

相比之下，大样本的定量分析却保留了大多数实验性的关注点，即利用跨个案分析来估计排除了其他所有影响的净因果效应。当然，时间序列统计技术——与过程追踪法自然对等的定量方法——已经存在了几十年（如 Box and Jenkens，1976）。然而，纯粹的时间序列分析在政治科学和社会学中却很少见。间歇时间序列分析（Interrupted Time Series Analysis）也有很长的历史（如 Campbell and Stanley，1963），但是，随着最近人们关注断点回归设计（regression discontinuity designs）（如 Dunning，2012），它获得了新生和新名。在实践中，统计结果常常依赖于时间序列和截面方差（cross-sectional variance）的（几乎未被分析过的）各类混合体。对于任何给定的研究来说，人们可以询问，方差中有多少是截面的，又有多少是时间序列的。然而，甚至在某些看似有很大时间序列成分的领域（例如，在 30 年时间里对 20 个工业国进行比较）中，大多数变异也都是截面的。因此，虽然许多定量分析考察了一个个案随着时间的变化过程，但是截面因素则完全驱动了坊间的研究成果。

7.3 因果过程观测比对数据集合观测

在两种传承中，个案内对比跨个案分析哪个相对重要，这个问题在科利尔、布雷迪和西赖特（Collier，Brady and Seawright，2010）讨论主

要的观测类型时被强调过。他们强调，这一点是在定性对比定量研究中获得因果推断的手段。他们将定性研究中的因果推断和“因果-过程观测”(causal-process observations, CPOs)联系起来，这意味着并要求有个案内分析。相比之下，在定量研究中，因果推断所采用的主要观测是“数据集合观测”(date-set observations, DSOs)，这种观测预设了一个截面研究设计。

因果过程观测项(CPO)是“一种洞见或一组数据，它提供了有关情境或机制的信息，并在因果推断中提供手段”(Collier, Brady and Seawright, 2010:184)。正如这个概念的标签和定义所表明的，因果过程观测是从个案中获得的一组特殊信息，这些信息允许研究者去评估某个给定的原因项是否被一个假设或理论赋予了要承担的原因角色。

下面三个例子有助于说明因果过程观测在定性个案内分析中的应用：

1. 吕贝特(Luebbert, 1991, 308—309)利用来自德国内部的因果过程观测项(CPO)来驳斥格申克龙(Gerschenkron, 1943)和穆尔(Moore, 1966)的命题，即法西斯政权是由一群地主精英导致的，这些精英促使众多农民都支持法西斯主义。他呈现的证据表明，德国的地主精英并没有得到大量农民的选票。实际上，农村是支持法西斯主义的，但这主要出现在农民家庭占主导的地区，而不是精英所居住的地区。这些观测尽管数量小而且只来自一个单独的个案，却有力地挑战了格申克龙-穆尔关于法西斯主义起源的命题。

2. 沃尔特(Walt)提出的假设是：革命引起战争。为了检验这个假设，他提出了一种理论来解释革命通过哪些干涉机制促成战争。这些机制包括革命性国家的偏好和生产力的变化，以及革命性国家和其他国家的关系的变化(比如，制造新的利益纷争和螺旋上升的怀疑)。如果该理论是正确的，这些机制就会表明特殊的因果过程观测(CPO)会出现在个案中。沃尔特的经验分析关注七项个案研究，其中革命都和战争相关联。然而，该论点的优势并不是小样本中的革命和战争之间有简单相关关系。毋宁说，对该论点的经验支持主要源自如下事实，即沃尔特在每一个个案中都能指向因果过程观测，这种因果观测对应于

他预测的机制。

3. 坦嫩瓦尔德(Tannenwald, 1999)认为，“核禁忌”——一种抵制核武器的规范禁令，它“将核武器指责为一种不可接受的大规模杀伤性武器”——是美国自第二次世界大战以来不使用核武器的原因。贝克(Beck, 2006)提出这样的疑虑：坦嫩瓦尔德只分析了四个数据集合观测值(DSO)(即四个历史阶段)，样本太小。相比之下，布雷迪、科利尔和西赖特(Brady, Coiller and Seawright, 2006)指出，在坦嫩瓦尔德的研究中进行因果推断的主要手段来自因果过程观测(CPO)，而不是数据集合观测(DSO)。尤其是，坦嫩瓦尔德呼吁重视高层决策者之间的特殊对话，这些对话表明对核武器的持续讨论甚至考虑都被主流规范所抑制。她的研究的优势并不建立在主要对四个历史时期的比较上，而建立在她有能力找到重要的个案内证据，这些证据表明，在有关核武器使用的争论中有越来越多强力的规范禁令。

正如这些例子所表明的，定性研究者把因果过程观测和对个案中发生的事件和过程的研究结合起来。在这些研究中收集的因果过程观测项并不是可以被聚合到一个矩形数据集合中的变量值。它们是特殊的个案内观测项，承载着针对该特殊个案要考量的假设。例如，坦嫩瓦尔德的一些关键数据就是来自外交政策制定者的一些特殊陈述，他们在其能力范围内好像被迫去提高使用核武器的可能性。她的论点的力度源自如下事实，即在美国除非有一些规范禁令使核武器的使用污名化了——即核禁忌，否则这些论述都将不复存在。同样，吕贝特关于德国的专业知识也为他提供了专业信息，这些信息严肃地挑战了这样的观念，即地主精英要为投票支持法西斯负责。只表面了解德国历史的学者是不能给出这种论证的。

相比之下，定量研究中的标准观测是数据集合观测，它相当于矩形数据集的行，即给定的某个个案在所有被测量变量上的数值。在主流的统计研究中，增加数据集合观测是提高自由度的一个标准方法。来自特定个案内部的潜在孤立和不可比较的观测——即因果过程观测(CPO)——并不用在定量范式中。如果信息只适用于一个或少数个案，这样的信息会被抛弃，因为它不能和统计检验联合起来应用。

在为了统计检验而收集数据集合观测项的时候要求有个案内数据，但这只在如下意义上才有如此要求，即分析者必须通过大量个案来测量特定的变量。分析者无须持有关于任何特殊个案的专门知识。个体个案的历史细节及特殊性质与统计检验无关。原则上讲，对于个体个案来说，人们除了知道他们在所测量的变量上的值之外，可以在对他们的诸多特征知之甚少的情况下很好地检验一个因果模型。①

总之，当定性和定量分析者在他们的研究中"加入新的观测"的时候，它们通常意味着不同的事情。对于定性研究者来说，这通常意味着在某些特殊的个案内部发现了新的证据或事实。这类似于在侦探工作中发现新的线索：新事实被揭露出来，让人们可以在与特殊的个案相关的假设和理论方面做出强有力的推断。相比之下，对于定量研究者而言，加入新观测项通常意味着加入了新的个案，即加入了一些主要分析单位的例子。在定量传统中，难以想像加入了观测却没有增加样本规模。但是在定性研究中，加入因果过程观测项通常不影响个案数量。实际上，在个案研究中，人们可以只研究一个个案（即一个数据集合观测项），但是有多个有效的因果过程观测项来为因果推断提供一个有力的支持（Collier, Brady and Seawright, 2010）。

7.4 过程追踪检验对比统计检验

过程追踪法在定性研究中得到了广泛使用，这也表明了个案内分析对于定性传统的重要性。尽管对于定性研究者来说，过程追踪并不是唯一可行的个案内模式的假设检验（参见第 9 章），但它是一种处于中心地位的定性方法，它与基于跨个案分析的假设检验的统计方法形成了鲜明的对比。

① 然而，要想正确地设定一个能够得到统计检验的因果模型，往往要求对个案的性质有很好的了解。

在一个特殊个案中的特殊结果有哪些原因，人们可就此提出一些假设，过程追踪法被用来评估此类假设。①这种方法建立在两种主要的检验基础上：箍筛检验和确凿证据检验（Van Evera，1997）。正如我们将看到的那样，在过程追踪检验、关于必要或充分条件的概括和特定个案内观测（即因果过程观测）的应用之间存在着固有的联系。

7.4.1 箍筛检验

箍筛检验指的是，要想保证关于单个个案的假设是有效的，某项给定的证据必须存在于该个案中（Van Evera，1997）。虽然通过了箍筛检验并不会确证一个假设，但是不通过箍筛检验却能拒绝假设。在此意义上，箍筛检验提供的证据是使假设有效的一个必要条件（Bennett，2008）。

箍筛检验可以涉及自变量、因变量和机制（参见 Collier，2011；Mahoney，2010，forthcoming），当该检验涉及自变量和因变量的时候，分析者需要用因果过程观测来确定的是，原因和结果的发生是否遵循了由所探究的假设来设定的方式。箍筛检验通常通过质疑一个个案的描述性事实来挑战一些假设。

例如，如第 4 章所讨论，在 2000 年佛罗里达州的总统选举中，媒体提前宣布了阿尔·戈尔的胜利，布雷迪（Brady，2010）在分析这种提前宣布的影响时使用了箍筛检验。他在使用箍筛检验过程中要识别一些必要条件，这些条件令媒体宣告造成布什损失掉佛罗里达州选民 i 的选票。这些必要条件包括以下四种：(1)选民住在佛罗里达州东部的一些走廊县；(2)媒体宣布时，该选民还没有投票；(3)选民能听到媒体宣布，并且(4)选民支持布什。

在这些箍筛检验的基础上，布雷迪考察了可能不投票给布什的选民，他排除了几乎所有的佛罗里达选民。实际上，他估计只有 280 位选民能通过全部四个检验。他认为在这 280 位选民中，大多数人没有被阻止去投票（比如，人们出于多种原因才投票，而不出于如下事实，即他

① 过程追踪分析也可以被应用于假设和理论的公式中。

们的单张选票将决定总统人选)。这使得他的估计值降低到28—56张缺失的选票(他将估计值乘以4,达到上线估计值224)。但是,正是箍筛检验才在布雷迪的论证中担当重任。

实际上,分析者也可以用箍筛检验去探究在一个自变量和一个因变量之间是否有因果联系。这些箍筛检验将注意力引向在设定好的原因和结果之间的干涉机制。例如,考虑一下前文提到的吕贝特反驳格申克龙-穆尔的假设(即在德国,压榨劳动力的地主精英是法西斯出现的一个关键原因)。要想使该假设有效,一个必要条件是地主精英直接或间接地控制农民,使农民投票支持法西斯主义候选人。然而,吕贝特利用来自德国的个案内数据表明这种机制是不存在的。德国的地主精英要么不会拉农民的选票,要么如果可以的话,他们更关心的是如何维持庇护网并实际支持自由党候选人。因此,该假设没有通过箍筛检验,只能放弃掉。

没有通过箍筛检验可以视为证伪一个假设的标准方法。但是,通过了箍筛检验就能为支持一个假设提供强有力的证据吗?答案是"取决于具体情况"。尤其是,这取决于箍筛检验的相对难度。通过了一个困难的箍筛检验的确能为一个假设提供实质性的积极支持,但通过一个简单的箍筛检验却不能。就像某些圆箍因比其他圆箍小而更难以跨越一样,一些箍筛检验的要求更高,因而难以通过(Mahoney, forthcoming)。

一条假设的成立需要有必要条件,通过一个箍筛检验的相关难度与这样的条件在数据或现实中出现的频次直接相关。如果这些条件总是存在,箍筛检验就容易,因为假设将几乎会自动通过检验。相比之下,如果使假设得以成立的必要条件在给定的环境中很少见或异常,箍筛检验就很难通过(参见 Hart and Honoré, 1985; Braumoeller and Goertz, 2000; Goertz, 2006; Ragin, 2008)。

7.4.2 确凿证据检验

确凿证据检验是这样的:如果存在某些特殊的证据(即特殊的因果过程观测项),假设就必然成立。通过了确凿证据检验,就会为假设提

供决定性的支持，尽管没有通过确凿证据检验也不会消除该假设。在此意义上，在确凿证据检验中被确认的因果过程观测项的存在就是使该假设成立的一个充分条件。①

和箍筛检验一样，确凿证据检验也涉及一个自变量、一个因变量或一种机制。例如，在坦嫩瓦尔德的研究中，她用确凿证据检验确立了她的主要自变量——核禁忌——的存在。她有一些特殊的例子，其中的决策者在关于外交政策的讨论中，通常在其能力范围内恰好被限制提出使用核武器的问题。核心的确凿证据是决策者的闭门讨论。如果没有核禁忌形式的证据(gun)，他们的评论——即确凿(smoke)——也就没有任何意义了。

虽然通过确凿证据的检验可以作为证据有力地支持一个假设，但是没有通过一种确凿证据检验的结局就各有不同了。某些确凿证据检验比其他检验更不容易通过。没有通过一种容易的确凿证据检验，就证明这个假设无效。例如，现在考虑一下约翰·斯诺(Snow, 1855, 1965；也见 Freedman, 1991)的著名研究，它证明，霍乱是一种传染性的疾病，而不是空气中有害气体(即瘴气)的产物。要给出支持瘴气理论的一个相对容易的确凿证据检验，就需要表明疾病有时候即使不通过人类携带的方式，也会传播到其他地方(如岛屿)。然而斯诺的发现恰恰相反：疾病总是跟随人类的旅途。尽管这个证据还没有完全反驳瘴气理论，但它当然是一重击。相比之下，即便不通过一个困难的确凿证据检验，也不会提供太有反驳性的证据。如果确凿证据检验相当难以通过，即便一个假设是有效的，人们也未必期待它得到通过。例如，在佛罗里达的走廊地区，有 1 万名特定的选民要投票给布什，但是当他们听到了媒体先行公布的选举结果后，决定不投了。具体地确认这 1 万名选民，就成为支持洛特(Lott, 2000)观点的确凿证据。然而，这是一个非常困难的确凿证据检验，事实上它不会得到支持，这个事实不会对洛特的观点产生多大冲击。相反，布雷迪(Brady, 2010)利用箍筛检

① 舍洛克·霍姆斯(Sherlock Holmes)有一句名言：一旦你消除了所有不可能的解释(即箍筛检验)，剩下的一个必然为真。这表明，人们完全用箍筛检验就能获得正确的解释。

验来反驳这个观点。

一项确凿证据检验的相对难度是与检验中使用的条件的普遍共性(即因果过程观测)相关的。所有的确凿证据检验都要参考这样一个条件(或者是条件的组合),即该条件的存在足以使所研究的假设成立。然而,这个条件出现的频次各不相同。在一种已知的情景下进行检验时,如果条件经常存在或"正常",但是假设可能没有通过确凿证据的检验,那么,这样的假设比那些条件很少出现情境下没有通过确凿证据检验的假设更可能是错的。

在社会科学中,我们怀疑学者通常不能使用确凿证据检验来确证假设。收集到的典型证据更像是弹壳(shell castings)而不是确凿证据:它的出现倒是一项确凿证据,但是确凿证据本身却没有被观测到。通常情况下,需要将一些关键证据组合起来形成一个真正使人信服的证据。继续这个隐喻,如果事实是嫌疑犯有一把枪与那些弹壳相匹配,再将出现的弹壳和这个事实结合,这样的话,弹壳看起来就更容易作为证据去确证这个假设。或者我们回到斯诺对霍乱的研究:他用多种不同的证据去支持传染病理论(参见第 8 章)。

定性研究者经常将多种证据组合起来以尝试获得充分性,这恰好符合在该传统中应用的典型的因果模型——即它是一种组合模型,其中的因素组合对于结果来说是充分的。结局就是,对一个个案的令人信服的解释通常要求有多条证据,这些证据加在一起构成了一个确凿证据的解释。单独一项观测不足以支持一个假设。相比之下,多条个别的证据很容易意味着箍筛检验的失败,因此拒绝一个假设。这有助于解释为什么定性研究者经常发现很容易使用个案内分析来消除一个假设,而不是确信地支持一个假设。

7.5 结论

跨个案分析和个案内分析分别是大样本分析和小样本分析中的核

心的因果推断模式，虽然如此，人们当然没有理由不能设计出一项诉诸这两种分析方式的研究计划。事实上，在两种传承中，最好的实践所指向的研究设计通常将二者组合在一起。

如在下一章中会进一步讨论的那样，在定量方面，统计研究者会越来越普遍地用定性的个案研究来补充他们的研究工作。来自回归分析的一个原因变量可能用个案研究加以考察，以便确定它是否按照正在检验的理论所设定的方式起作用。在这种补充性的分析中，研究者使用个案内分析，同时搜索这样的一些因果过程观测项，即这些观测项在评估一个变量的因果影响方面具有可提供证据的价值。在定性研究方面，当研究者就一个或少量个案做出一个发现的时候，人们不禁会问，这个发现是否具有普遍意义（参见第 16 章）。因此，小样本的发现可能会刺激一个范围更广的利用大样本的跨个案检验（Lijphart，1971）。

总之，大多数实际的研究主要依赖于跨个案分析或个案内分析，尽管如此，这两种模式能够并且（根据许多重要的方法论学者的看法）应该经常被组合起来。不同类型的多元方法研究越来越流行，这表明学者正愈加寻找一些方式来达到这种综合。

参考文献与推荐读物

对个案内分析的讨论（在各种名义下）在定性方法论领域已经有很长历史。参见巴顿和拉扎斯菲尔德（Barton and Lazarsfeld，1955）的“过程分析”（process analysis），斯梅尔瑟（Smelser，1968：72—73；1976：217—218）的“单位内”（intra-unit）或“单位内比较”（within-unit comparison）以及坎贝尔（Campbell，1975：181—182）的“模式匹配”（pattern matching）。关于过程追踪的经典研究是乔治和麦基翁（George and McKeown，1985）；当代对它的一个重要表述是乔治和贝内特（George and Bennett，2005）。也可参见贝内特（Bennett，2006，2008）给出的重要推进。范·埃弗拉（Van Evera，1997），科利尔（Collier，2011）和马奥尼（Mahoney，forthcoming）讨论了箍筛检验、确凿证据检验和迹象检验（straw in the wind tests）。因果过程观测和数据集合观测之间的差异在科利尔、布雷迪和西赖特（Collier，Brady and Sea-

wright, 2010)的研究中得到了推进。有关不同类型的因果过程观测的研究,参见马奥尼(Mahoney 2010)。

佩弗豪斯和布罗泽克(Pevehouse and Brozek, 2008),贝克(Beck, 2008)述评了关于时间序列分析和时间序列截面分析(time-series cross-sectional analysis)的文献。人们尝试将中介过程整合进入潜在结果统计架构之中,有关这方面的精致尝试,参见格林和奎因(Glynn and Quinn, 2011),今井、基尔、廷利和山本(Imai, Keele, Tingley and Yamamoto, 2011)。也可参见珀尔(Pearl, 2010),格林、尚和布洛克(Green, Shang and Bullock, 2010)。

Barton, A. H., and P. F. Lazarfeld. 1955. "Some Functions of Qualitative Analysis in Social Research." *Sociologica* 1:324—361.

Beck, N. 2006. "Is Causal-process Observation an Oxymoron?" *Political Analysis* 14:347—352.

Beck, N. 2008. "Time-series Cross-sectional Analysis." In *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmeier, H. Brady, and D. Collier. Oxford: Oxford University Press.

Bennett, A. 2006. "Stirring the Frequentist Pot with a Dash of Bayes." *Political Analysis* 3:339—344.

Bennett, A. 2008. "Process Tracing: a Bayesian Perspective." In *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmeier, H. Brady, and D. Collier. Oxford: Oxford University Press.

Bennett, A., and C. Elman. 2006. "Complex Causal Relations and Case Study Methods: the Example of Path Dependence." *Political Analysis* 14:250—267.

Blalock, Jr., H.M. 1964. *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

Bollen, K.A. 1989. *Structural Equations With Latent Variables*. New York: Wiley.

Box, G., and G. Jenkins. 1976. *Time Series Analysis: Forecasting and Control*, revised edition. San Francisco: Holden Day.

Brady, H.E. 2010. "Data-set Observations Versus Causal-process Observations: the 2000 U.S. Presidential Election." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E. Brady and D.Collier. Lanham, MD: Rowman & Littlefield.

Brady, H.E., D.Collier, and J. Seawright. 2006. "Toward a Pluralistic Vision of Methodology." *Political Analysis* 14:353—368.

Braumoeller, B.F., and G. Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44: 844—858.

Campbell, D.T. 1975. "'Degress of Freedom' and the Case Study." *Comparative Political Studies* 8:178—193.

Campbell, D. T., and J. C. Stanley. 1963. *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin Company.

Collier, D. 2011. "Understanding Process Tracing." *PS: Political Science and Politics* 44:823—830.

Collier, D., H. Brady, and J. Seawright. 2010. "Sources of Leverage in Causal Inference: toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E.Brady and D.Collier. Lanham, MD: Rowman & Littlefield.

Collier, R.B, and D.Collier. 1991. *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America*. Princeton: Princeton University Press.

Dunning, T. 2012. *Natural Experiments in the Social Sciences*. Cambridge: Cambridge University Press.

Freedman, D. 1991. "Statistical Models and Shoe Leather." In *Sociological Methodology*, edited by P. Marsden. San Francisco:

Jossey-Bass.

George, A., and A.Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MIT Press.

George, A.L., and T.J. McKeown 1985. "Case Studies and Theories of Organizational Decision Making." In *Advances in Information Processing in Organizations*, vol.2, edited by R.F.Coulam and R.A. Smith. Greenwich, CT: JAI Press.

Gerschenkron, A. 1943. *Bread and Democracy in Germany*. Berkeley: University of California Press.

Goertz, G. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41: 88—109.

Glynn, A.N., and K.M., Quinn. 2011. "Why Process Matters for Causal Inference." *Political Analysis* 19:273—286.

Green, D.P., E.H. Shang, and J.G. Bullock. 2010. "Enough already about 'Black Box' Experiments: Studying Mediation is more Difficult than most Scholars Suppose." *Annals of the American Academy of Political and Social Science* 628:200—208.

Hart, H.L.A., and T.Honoré 1985. *Causation in the Law*, 2nd edition. Oxford: Oxford University Press.

Imai, K., L.Keele, D.Tingley, and T.Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105:765—789.

Lijphart, A. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65:682—693.

Luebbert, G.M. 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. Oxford: Oxford University Press.

Mahoney, J. 2010. "After KKV: the New Methodology of Qualitative Research." *World Politics* 62:120—147.

Mahoney, J. Forthcoming. "The Logic of Process Tracing in the Social Sciences." *Sociological Methods and Research*.

Moore, Jr., B. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Pearl, J. 2010. "The Mediation Formula: a Guide to the Assessment of Causal Pathways in Non-linear Models." Technical Report R-363.

Pevehouse, J.C., and J.D. Brozek. 2008. "Time-series Analysis." In *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmeier, H. Brady, and D. Collier. Oxford: Oxford University Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Smelser, N.J. 1968. *Essays in Sociological Explanation*. Englewood Cliffs, NJ: Prentice-Hall.

Snow, J. 1855, 1965. *On the Mode of Communication of Cholera*, 2nd edition. London: John Churchill.

Tannenwald, N. 1999. "The Nuclear Taboo: the United States and the Normative Basis of Nuclear Non-use." *International Organization* 53:433—468.

Van Evera, S. 1997. *Guide to Methods for Students of Political Science*. Ithaca: Cornell University Press.

Walt, S.M. 1996. *Revolution and War*. Ithaca: Cornell University Press.

8 因果机制与过程追踪

当运用过程追踪法进行因果推断的时候，通过引用因果机制来解释便……提供了因果推断的有力来源，它可以在一些单独个案内细致地考察过程。

——亚历山大·乔治(Alexander L.George)
安德鲁·贝内特(Andrew Bennett)

8.1 导言

我们对因果性的直觉理解包含这样一个生成过程，在这个过程中，一个原因通过引发某些机制和过程的运行而导致一个结果。当人们面对的数据表明两个变量之间有关联的时候，他们在声称这种关联本质上是因果关系之前，通常还额外需要与机制有关的信息(Ahn, Kalish, Medin and Gelman, 1995)。社会科学家也不例外：他们相信因果效应要通过一个过程传递到另一个过程。

在社会科学和科学哲学领域，围绕“因果机制”观念提出了大量的文献，其中包含一系列异质的观点和定义(参见本章给出的推荐读物)。基于我们的目的，这里无须钻研其细节。相反，我们可以这样来理解因果机制，即它是原因导致结果所历经的介入过程。我们认为，任何得到相对较好发展的理论都会提供有关因果机制的讨论。这同样适用于在定量和定性研究传统中检验了的理论：这些理论提出的因果机制观点

将自变量和因变量联系起来。

本章探讨的关键问题是，当定性和定量传统在做因果推断时，如何经验地评估有关机制的理论。在定性传承中，研究者在评估时，尝试通过过程追踪和对因果过程观测的分析来观测机制（Collier, Brady and Seawright, 2010a, 2010b; George and Bennett, 2005）。在定性范式中，对特殊个案的个案内分析很自然地和观测机制的尝试结合在一起。

相比之下，统计方法的设计可不是为了在特殊个案内部观测机制。使用统计量来推断——不管是基于观测数据还是实验数据——都依赖于对多项观测的跨个案分析。在这一传统中，如果一个变量在一次统计检验中发挥了所预测的效应，那么研究者可能预期某种给定的机制起了作用。然而，他们通常不会从经验上去研究机制自身。变量为什么产生作用，这是理论的一部分，但是它通常不包含在统计检验中（参见附录）。

8.2 机制和因果推断

人们早就通过一些伪相关的例子（例如，一个地区的麻雀数量和该地区的生育率之间有关联）得出"相关不是因果"的结论。学生在第一门方法课上，就被教导去想象有没有第三个先导变量可以引起这两个变量，并因此可以解释相关（例如，一个测量城乡位置的变量可能既解释麻雀数量，也解释生育率）。然而，当同学们第一次遇到这些例子的时候，他们怀疑这种相关可能是伪相关，其原因通常在于这里缺乏直觉的因果机制，而不是因为他们立即意识到存在一个特定的先导变量可以解释掉这个相关。人们之所以怀疑麻雀-生育率相关问题不是因果关系，是因为这里没有看似可行的机制（Porpora, 2008）。

在社会科学中，许多统计方法论学者都认定对观测数据进行因果推断是极其困难的。观测性研究缺乏受控实验的随机安排，要求用控

制变量来处理混淆因素。我们常常从年轻的定量方法论学者那里得知，回归无非是数据描述。①它最多会被认为是一种生硬的因果推断工具（参见 Collier, Brady and Seawright, 2010b）。

可见，用观测数据做因果推断是有挑战性的，学者已经认识到这一点，进而激起了在实验方面逐渐增长的兴趣。社会科学现在从事着各类实验研究，包括调查实验、实验室实验和田野实验（当然，心理学早就一直在使用实验）。在政治科学顶尖杂志中常常看到利用实验法的论文。甚至当方法论者讨论回归设计时，现在也经常接受一些实验术语，如处理和控制。实际上，对于某些定量方法论学者而言，新的口号也许是：

没有实验，就得不出强有力的因果推断。

有了良好的实验设计，人们就可以在没有观测到因果机制的情况下来评估一个给定处理的平均效应。正如格林及其同事所指出："仅通过操控变量并测量其对结果的影响都可以了解很多理论的和实际的价值，不管这些影响通过怎样的因果途径。"（Green, et al., 2010:206—207）②然而，如同在所有的科学中一样，在社会科学中，如果可能的话，学者仍然希望填满实验的黑箱。当研究者呈现实验结果的时候，他们通常必须回答有关机制的问题，这种机制把处理和效应联系在一起。他们尝试努力做出解答，因为一个良好的理论会在一个观测到的效应背后确立其机制。

相比之下，在定性研究中，研究者将机制的鉴别当做因果推断的关键。他们把机制看做一种非实验性的、将因果关系从伪相关中区分出来的方式。

① "如果没有实验、自然实验、中断或其他强有力的设计，再多的计量经济学或统计建模也不能令人信服地从相关走向因果。"（Sekhon, 2009:503）

② 用实验法来研究中介项，这项事业面临着很多困难（参见 Bullock and Ha, forthcoming）。此外，用观测数据对中介项进行评估的统计技术要求有非常强的假定，因此在实践中难以实施（例如，Imai et al., 2011）。

> 在因果推断中，机制可以通过两种方式起作用。X 可以通过某种机制影响 Y，认识到这种机制的存在即可支持一个推断，即 X 是 Y 的一个原因。此外，如果缺乏将 X 和 Y 联系起来的看似可行的机制，我们就有很好的理由去怀疑二者之间是否有因果关系。……尽管这种怀疑可能太强了，以至于不能说机制的设定对于因果推断来说总是必要的，然而一个令人完全满意的社会科学解释还是要求对这些因果机制加以设定。(Hedström and Ylikoski, 2010:54；也可参见 George and Bennett, 2005)

人们甚至可能认为在定性传承中已经发展出了一种规范，它使强因果推断要求在个体个案内部进行过程追踪，以便考察所提出的因果机制是否存在。如此看来，对于定性学者来说，口号或许是：

> 没有过程追踪，何谈强因果推断。

因此，定量和定性范式关于强因果推断有不同的看法。毫不奇怪的是，一方可能对另一方的标准抱有怀疑。例如，过程追踪可以为因果推断提供有力的支持，这种观念在定量传承中并没有得到广泛的认同。金、基欧汉和维巴提议到，过程追踪“不那么容易产生强因果推断”，只能“提升描述性的概括，为因果推断提供一种途径”(King, Keohane and Verba, 1994:227—228)。其他学者则强调，因果机制不是在因果分析中解决根本问题的“奇迹创造者”(如 Gerring, 2010; Norkus, 2004)。从统计视角来看，关于因果机制的推断必须符合好的因果推断的要求，即因果机制适用于任何潜在的处理或变量。因果机制不要求对因果性有新的理解(King, Keohane and Verba, 1994)，尽管在估计因果机制的效应时所涉及的计量经济学难题有特殊性(例如，ManKinnon, 2008)。

过程追踪可以在多个方面与大样本分析有交集(Collier, Brady and Seawright, 2006, 2010c)。有时因果机制首先在个案分析中被挖掘出来，然后利用大样本分析去确证这个发现。例如，可以考虑斯诺

的研究，他向我们展示了水（而非空气中的沼气）是霍乱传播的机制(Snow, 1855, 1965，也可参见 Freedman, 1991)①。斯诺像一位典型的定性分析者那样开始自己的研究：始于深入考察 $Y = 1$ 的个案，即霍乱病患者。他注意到，某种病原菌(causal agent)看起来首先攻击了消化道。这使受污染的水或食物可能成为疾病传递的机制。他进行了其他关键的观测：海员只有在靠岸或补给时才会感染疾病，这些疾病跟随商业路线传递，而如果个体居住的房屋有单独的水供应，他们通常就不会感染疾病。他利用相邻的两个房屋设计了一种求异法：其中一个房屋有污染的水。他针对选出来的个体进行了同样的实验。然后，他用一种准实验令人信服地证明了该假设，这个准实验所用的数据来自大量的从不同渠道获取水的房间(Dunning, 2008)。大样本的自然实验确证了他此前通过定性研究得出的因果机制。

有时候，我们在观测数据中通常有一个大样本的统计发现，但是其因果机制有争议。在这种情况下，过程追踪也被用来对相互竞争的机制加以裁断。例如，一项长期的跨国定量研究发现，经济发展（经常用人均 GDP 来测量）和民主正相关（参见 Robinson, 2006 的述评）。实际上，这种关系在政治学和政治社会学中被认为是最稳健的发现之一（如 Geddes, 1999）。然而，几乎对于任何人来说，这项发现看起来都是未完成的，因为它留下了一个黑箱，而且不允许学者去评估其他备选的机制理论。对于定性研究者来说，必须用详尽的分析来充盈黑箱，即分析在一些特殊个案中导致民主的实际序列。为了区分出在因果关系被建立之前的机制，人们必须走出统计相关，走向定性分析。②

吕舍迈耶、斯蒂芬斯和斯蒂芬(Rueschemeyer, Stephens and Stephen, 1992)的《资本主义发展和民主》就是一个优秀的例子，该书尝试

① 这个例子是科利尔、布雷迪和西赖特(Collier, Brady and Seawright, 2006, 2010c)与贝克(Beck, 2006, 2010)之间争论的核心。

② 如吕舍迈耶、斯蒂芬和斯蒂芬斯所说："（发展和民主之间关系的）重复性的统计发现有一个特殊的'黑箱'特征，这个黑箱只有通过基于良好理论的经验分析才能被克服，……我们认为，比较历史研究最有可能点亮黑箱。……历史研究向事件序列及其与周遭结构条件的关系赋予洞见，在提供有效的因果解释时，这一点不可或缺。因果分析实质上是序列分析。"

利用历史研究，在个案内观测因果机制。他们提议到，发展促进了不同阶级（尤其是地主和工人）之间权力平衡的变化，而这种变化是民主的关键性机制。具体而言，他们给出的假设认为，发展可以促进"完全"民主的两个必要条件：(1)有权势的地主消失，(2)出现强力的、赞成民主的工人阶级。[①]尽管这些要素几乎是普遍的机制，他们却不是充分条件。民主化取决于与国家、政党、国际体制有关的其他要素。

在定性研究中，由于过程追踪是如此居于因果推断的核心，以至于该传统中的研究者会怀疑那些不能确认或检验因果机制的研究。例如，如果大样本的发现未经后续真正因果性过程追踪的证实，他们可能就不确信这些发现。通过细致考察个体个案，可能引发对来自统计学或形式分析的假定性因果机制的怀疑，这样的个案并不难发现。

1. 丘萨克、艾弗森和索斯凯斯(Cusack, Iverson and Soskice, 2007)认为，在西欧，商业和劳动力的经济倾向是将整合的劳动力市场和比例代表选举制(proportional representation electoral systems)的创立连接起来的关键机制。这些学者在劳动力市场整合和比例代表制之间找到了显著的统计关系。然而，他们并没有考察商业和劳动力的制度性倾向，这些倾向被假设会驱动这种关系。为了审查这种论证，克勒泽(Kreuzer, 2010)考察了历史研究是否提供了任何证据证明这些行动者关心选举制度的形式。在看完丘萨克、艾弗森和索斯凯斯的18个个案中的每一个之后，他总结道，他们提议的机制并不起作用："我不能找到任何证据将商业团体、工会、党派或者它们各自领导者的制度化倾向与劳动力市场关联起来。事实上，我不能发现任何证据表明商业或工会明确地偏爱某种选举制度而不是另一种。关于党派的制度化倾向已经有

① "民主只有在以下条件下才会建立起来：(1)地主的力量不强，或者(2)他们不依赖于大量廉价劳动力的供应，或者(3)他们不会控制国家"(Rueschemeyer, Stephen and Stepuhens, 1992:270)。"有组织的工人阶级在几乎任何地方都以一个关键行动者的身份出现在完全民主发展的国家中，唯一的例外是一些小型国家中的土地民主案例(agrarian democracy)。"(Rueschemeyer, Stephen and Stepuhens, 1992:270)

大量的讨论，但是没有任何一个指向经济因素。”(Kreuzer, 2010: 376)①

2. 科利尔和赫夫勒(Collier and Hoeffler, 2001)，费伦和莱廷(Fearon and Laitin, 2003)使用跨国统计分析，发现人均GDP与内战有强的负相关关系。然而，这两组作者在因果机制上出现分歧：科利尔和赫夫勒从贫穷对经济机遇的影响角度来理解机制，而费伦和莱廷却从国家避免内战的能力角度看待机制。基于个案研究的证据，桑班尼斯的发现只部分支持了二者提出的机制。他认为，人均GDP可能在与其他变量的交互作用中发挥影响：“不同的国家对待内战有不同的倾向，原因可能与其他自变量(如种族与民主)在各种收入水平上的表现方式有关。”(Sambanis, 2004:266)桑班尼斯表明，在科利尔、赫夫勒和费伦、莱廷的理论中提出的机制缺乏经验支持，令人质疑其实际效用。“如果大样本研究在因果路径方面做出了错误的假定，它们就会缺乏解释力……我们知道，通过增加人均GDP，就会在一定程度上降低内战的风险，但是一项更有针对性的政策干预可能更有效率，也更易执行。”(Sambanis, 2004:273)

3. 阿西莫格鲁和罗宾森(Acemoglu and Robinson, 2006)认为，在经济高度不平等的国家中更容易发生军事政变。他们运用博弈论区分出一种机制来解释这种关系：民主政治下的再分配数量在不平等的社会中会较高，因此，精英们会有更大的热情去招募军队推翻民主政治。然而，斯莱特和史密斯(Slater and Smith, 2010)批判了这种解释，他们的个案研究证据表明：“军队事实上几乎不是代理人，并且很少成为经济精英的同盟。”通常情况下，军队发生政变的原因与精英阶级的特殊经济利益无关。因此，他们认为，与阿西莫格鲁和罗宾森的博弈论模型相联系的因果机制在绝

① 丘萨克、艾弗森和索斯凯斯(Cusack, Iverson and Soskice, 2010)在反驳中认为，党的领导人参与经济事务，因而不希望工人及商业领袖积极推动一种特殊的选举制度。相反，他们认为党的偏好应当是在检验因果机制时所关注的核心。

大多数军事政变中并不存在。

总而言之，如果实验与非观测性的分析不能识别出机制，定性研究通常会质疑它们。在定性传承中，如果机制不为人所知，解释就是无力的。

8.3 多元方法中的过程追踪对比定性研究

随着多元方法研究的兴起，过程追踪(process tracing)不再是定性研究的特有领域了。在某些领域的定量研究者中间，过程追踪已成为一种"时尚需要"(de rigueur)，它在整体分析中包括单项个案研究。这种趋势与我们前文讨论的回归分析的滑坡有关。对所选的个案进行集中的过程追踪，认为它可以补充在不能进行实验的环境下的大样本研究。统计分析提供的证据可能表明，一个假想的因果机制在一个大的个案总体中起作用。然后，针对所选的个体个案进行过程追踪，即可用于探索因果机制是否如广而告之的那样起作用。这种多元方法策略突出地体现在最近的一些重要研究之中，这些研究首先给出大样本的统计结果，然后紧跟着对多个个案研究的分析(例如，Fortna，2007；Lange，2009；Lieberman，2003；Pevehouse，2005)。①

尽管人们趋于认为过程追踪可以作为一种有用的因果推断工具，不过，那些用过程追踪来补充统计发现的学者对该方法的使用好像与定性研究者的使用方式并不一样。过程追踪在补充统计发现中的运用因如下事实而变得复杂，即在统计分析中，随着从总体走向个案，因果过程未必稳定。例如，在一项统计研究中，设想一下 X_1 在总体中有显著的正影响。这是否意味着 X_1 不能对一个特殊的个案子集(subset

① 博弈论传统中的某些学者也转向了过程追踪，他们将其视为一种手段，用来检验他们的形式模型的可观测到的意义，例如贝茨等(Bates et al.，1998)。

of cases)发挥负影响？答案当然是“否”。在从超集到子集再到个案的过程中，X_1 的影响总是因统计模型而定；没有什么数学原因可以说明为什么 X_1 不能与特殊子集中的结果负相关。因此，当执行过程追踪时，我们不能确认因果机制在随意选择的特殊个案中是否如预期的那样是奏效的。①

为了用过程追踪来补充统计发现，分析者通常尝试挑选一个或多个个案，其中，所感兴趣的主要自变量应该扮演由理论所赋予的角色（参见第 14 章）。然而，当从总体走向子集合的时候，统计发现也可能不稳定，正因如此，分析者事实上从未在个体个案的数据集合值、统计模型中的参数估计值以及来自个体个案研究的观测值之间建立直接的关联。当统计研究者在特殊个案中运用过程追踪时，他们的特殊回归结果会趋于脱离图像。

始于统计模型的过程追踪研究者在进行个案研究时，通常将所感兴趣的自变量非正式地视为是一种有贡献的原因要素。在针对整个总体经常用到的加法模型(additive model)中存在着多种原因，但所关注的主要变量只是其中之一。可以将该变量的效应大致看成是对因变量的一种原因权重。因此，当在一个个案中针对该变量进行过程追踪时，研究者探索的是它如何为结果变量“做贡献”或“加权重”。然而，分析者通常并不认为个体变量对于结果来说是必要的。定性研究中的过程追踪是围绕反事实进行的，而当用过程追踪来补充统计学研究时却不如此。

当用过程追踪来补充统计分析时，它的另一个特点涉及其他变量（即控制变量，而不是所关注的变量）的角色。由于它们不是研究者特意关注的变量，在过程追踪分析中通常也不受重视。控制变量的作用必须被承认，但是注意力要转向感兴趣的主要变量。

在定性研究方面，研究者未必始于已经被过程追踪所验证的跨个案的发现。相反，他们可能从过程追踪开始，并用这种方法作为因果推

① 方法论者对此已有一个回应，即他们认真思考如何根据初步的回归结果来挑选用于过程追踪的个案（参见第 14 章）。

断的基础。然而，出于比较的目的，我们还是假定定性研究者始于一个运用于某个个案总体的集合论模型。该学者在进行个案内的过程追踪时，会针对因果关系保持一种集合论研究路径。这是对的，因为在集合论的架构下，某种因果模式可应用于作为一个整体的总体之中，而总体中的特殊个案会同样遵循该因果模式。当学者在执行过程追踪时，会认为个体个案中的原因对整个总体有相同的效应。

因果效应在从总体到个案中的这种稳定性最容易用必要条件来检视。在一个总体中，如果 A 是 Y 的一个必要条件，那么 A 必然是该总体中任意一个个案(或个案子集)的必要条件。就一个实质性的个案来说，现在考虑这样一个假设：威权政体是大屠杀的一个必要条件。如果该假设成立，则它会适用于任意一个大屠杀个案。在假定该假设一致地适用于所有大屠杀的情况下，人们就可以进行过程追踪了。

从总体到个案的这种稳定性也适用于分析诸如下面这样的一个集合论模型中的充分性：$Y = ABc + DE$。如果一个个案有某种组合(即或者是 $A * B * c$ 或者是 $D * E$)，它就会有感兴趣的结果。为了探究机制，过程追踪会选择这样的一个个案，其中或者出现 ABc，或者出现 DE，但两者不能都出现(参见第 14 章“个案选择与假设检验”)。然后，分析者会探索诸如 A、B 和 c 如何组合起来创造出结果。当进行过程追踪时，研究者将探究这种因果组合通过怎样的过程才产生了感兴趣的结果，要努力区分出这些特殊的过程。

当定性分析者就一个特殊个案中的某种因果组合进行过程追踪时(假定该个案只展现了可产生结果的一种组合)，他或她可以将该组合中的每一个自变量视为该个案有此结果的必要条件。也就是说，如果 $Y=1$ 这个个案只展示了一个因果组合，那么该组合中的每一个原因项对于在此个案中发生的结果来说都是必要的。如果模型正确，并且拥有 $D * E$ 组合的某个个案缺少了 D 或 E，那么该个案将会体验不到结果。因此，人们在针对一个因果组合中的个体变量进行过程追踪时，通常采用与必要条件相同的方式。

鉴于此，对于始于某种集合论模型的分析工作来说，当针对某个特定个案进行过程追踪时，它必须将被该个案拥有的因果组合中的每一

个变量都看成是必要的。然而，分析者通常不只关注一个特殊感兴趣的变量，因为该组合中的所有变量对于个案中的结果来说都是必要的。相比之下，当过程追踪与统计研究相结合时，关注点就是感兴趣的那个主要变量了。如果所关注的主效应是一个交互项，那么唯一用于补充统计研究的时间过程追踪法（time process tracing）就会关注一套原因（packages of causes）。然而，要想确证在某个统计模型的交互项背后有哪些假定的机制，分析者可以利用过程追踪法，不过这样的实例还不为我们所闻。

总之，当在个案中进行过程追踪时，每一种传承中的因果模型都倾向于保持其正确性。定性研究者运用的是一种集合论模型，该模型的基础在于，针对结果有诸多必要条件，还有合起来可充分带来结果的一系列条件。作为一般的规则，人们通过询问分析者是否将个体原因当做必要条件，并/或询问当他们讨论充分性的时候采用怎样的机制将这些条件组合在一起，这样便可以区分出一个过程追踪的定性研究。相比之下，在特定的个案内进行过程追踪时，始于统计模型的多元方法研究者通常接受一种加法的因果研究路数。他们通过过程追踪，探究感兴趣的单个因素是否为一个特定个案中的某个特定结果做出贡献或增加权重。由于有如此多的其他原因被假定是有作用的，因此，他们不再假定感兴趣的要素对于结果来说是必要的。

8.4 结论

在定性传承中，标准的和自然的做法是研究因果机制，并用过程追踪进行个案研究。通过在一个或多个特定的个案中追踪 X 导出 Y 的过程，就可以部分地推断 X 是 Y 的一个原因。该研究传统中的学者常采用个案内分析，这个事实也推动了过程追踪。这种分析模式本身旨在区分出诸多机制，一个特定的原因要素可以通过这些机制在一个特定的结果上产生其效应。

在定量传承中，人们越来越关注回归分析在产生强因果推断方面的能力，这一点在两个方向上推动了方法论进程。一个方向是，各类实验愈加显耀（尽管在所有研究中只占少数）。通过实验，人们可以在不检验机制的情况下，很好地估计某种处理的平均效应。然而，由于几乎所有的社会科学理论都倡导关于机制的思想，因此，可以认为实验研究留下的黑箱是有问题的。

另一个方向为，多元方法研究越来越被认为是最好的实践，在此类研究中，定量分析者将回归分析与个案研究分析结合起来。如果一个变量在回归分析中有显著效应，就会进一步在个案研究中考察它，这样即可决定它是否按照正在被检验的理论所假定的方式起作用。然而，与定性传统不同的是，多元方法研究中的过程追踪是依据加法模型来审视原因的。研究者不用过程追踪去检验 X 是否为 Y 的必要条件。相反，其主要目标在于探索 X 是否对 Y 的水平或出现有所贡献。

参考文献与推荐读物

有关因果机制的文献体现了对因果机制的多种不同的理解和界定。各种具体的观点，可参见邦奇（Bunge，1997）；埃尔斯特（Elster，1989）；法莱蒂和林奇（Falleti and Lynch，2009）；耶林（Gerring，2008）；赫德斯特伦和斯韦德贝格（Hedström and Swedberg，1998）；马奥尼（Mahoney，2001）；迈因茨（Mayntz，2004）；麦克亚当、塔罗和蒂利（McAdam，Tarrow and Tilly，2001）；诺尔库斯（Norkus，2004）；塞科恩（Sekhon，2004）；斯汀康比（Stinchcombe，1991）和瓦尔德纳（Waldner，2012）。有关过程追踪方法的文献，参见乔治和贝内特（George and Bennett，2005）；贝内特（Bennett，2006，2008）；霍尔（Hall，2003）；罗伯茨（Roberts，1996）；范·埃弗拉（Van Evera，1997）；科利尔（Collier，2011）和马奥尼（Mahoney，forthcoming）。科利尔、布雷迪和西赖特（Collier，Brady and Seawright，2006，2010a）提出了因果过程观测（causal-process observation）这个概念。有关多元方法研究的文献，参见耶林（Gerring，2011）和利伯曼（Lieberman，2005）。有关过程追踪法通过哪些方式可以补充集合论研究，参见施耐

德和罗尔夫林(Schneider and Rohlfing, 2011)。关于利用实验法和统计法来研究机制的文献,参见布洛克和哈(Bullock and Ha, forthcoming);格林和奎因(Glynn and Quinn, 2011);格林、哈和布洛克(Green, Ha and Bullock, 2010);霍兰(Holland, 1988);今井等(Imai et al., 2011)和麦金农(MacKinnon, 2008)。虽然实验法可以解决观测研究的某些问题,但是它们自己也有问题。有关实验法的前景和陷阱的比较好的讨论,参见德鲁克曼、格林、库克林斯基和卢皮亚(Druckman, Green, Kuklinski and Lupia, 2011),莫顿和威廉姆斯(Morton and Williams, 2010)。

Acemoglu, D., and J.A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Ahn, W., C.W.Kalish, D.L.Medin, and S.A.Gelman. 1995. "The Role of Covariation Versus Mechanism in Causal Attribution." *Cognition* 54:299—352.

Bates, R., A. Greif, M. Levi, J.-L.Rosenthal, and B.Weingast. 1998. *Analytic Narratives*. Princeton: Princeton University Press.

Beck, N. 2006. "Is Causal-process Observation an Oxymoron?" *Political Analysis* 14:347—352.

Beck, N. 2010. "Causal Process 'Observation': Oxymoron or (fine) Old Wine." *Political Analysis* 18:499—505.

Bennett, A. 2006. "Stirring the Frequentist Pot with a Dash of Bayes." *Political Analysis* 3:339—344.

Bennett, A. 2008. "Process Tracing: a Bayesian Perspective." In *The Oxford Handbook of Political Methodology*, edited by J.Box-Steffensmeier, H.Brady, and D.Collier. Oxford: Oxford University Press.

Brady, H.E. 2010. "Data-set Observations Versus Causal-process Observations: the 2000 Presidential Election." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by

H.E.Brady and D.Collier. Lanham, MD: Rowman Littlefield.

Bullock, J.G., and S. E.Ha. Forthcoming. "Mediation Analysis is Harder than it Looks." In *Cambridge Handbook of Experimental Political Science*, edited by J.Druckman, D.Green, J.Kuklinski, and A.Lupia. New York: Cambridge University Press.

Bunge, M. 1997. "Mechanism and Explanation." *Philosophy of the Social Sciences* 27:410—465.

Collier, D. 2011. "Understanding Process Tracing." *PS: Political Science and Politics* 44:823—830.

Collier, D., H.E. Brady, and J. Seawright. 2006. "Toward a Pluralistic Vision of Methodology." *Political Analysis* 14:353—368.

Collier, D., H.E. Brady, and J. Seawright. 2010a. "Sources of Leverage in Causal Inference: toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E.Brady and D.Collier. Lanham, MD: Rowman & Littlefield.

Collier, D., H.E. Brady, and J. Seawright. 2010b. "Introduction to the Second Edition: a Sea Change in Political Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edition, edited by H.E.Brady and D.Collier. Lanham, MD: Rowman & Littlefield.

Collier, D., H. E. Brady, and J. Seawright. 2010c. "Outdated Views of Qualitative Methods: Time to Move on." *Political Analysis* 18:506—513.

Collier, P., and A. Hoeffler. 2001. "Greed and Grievance in Civil War." World Bank Research Working Paper, No.2355.

Cusack, T., T. Iverson, and D. Soskice. 2007. "Economic Interests and the Origins of Electoral Systems." *American Political Science Review* 101:373—391.

Cusack, T., T. Iverson, and D. Soskice. 2010. "Coevolution and

Political Representation: the Choice of Electoral Systems." *American Political Science Review* 104:393—403.

Druckman, J.N., D.P.Green, J.H.Kuklinski, and A.Lupia, eds. 2011. *Cambridge Handbook of Experimental Political Science*. New York: Cambridge University Press.

Dunning, T. 2008. "Improving Causal Inference: Strengths and Limitations of Natural Experiments." *Political Research Quarterly* 61:282—293.

Elster, 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Falleti, T.G., and J.F.Lynch. 2009. "Context and causal mechanisms in political analysis." *Comparative Political Studies* 42: 1143—1166.

Fearon, J.D., and D.D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97:75—90.

Fortna, V. 2007. *Peacekeeping and the Peacekept: Maintaining Peace After Civil War*. Princeton: Princeton University Press.

Freedman, D.A. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology* 21:291—313.

Geddes, B. 1999. "What do we know about Democratization after Twenty Years?" *Annual Review of Political Science* 2:115—144.

George, A., and A.Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Gerring, J. 2008. "The Mechanistic Worldview: Thinking inside the Box." *British Journal of Political Science* 38:161—179.

Gerring, J. 2010. "Causal Mechanisms, yes, but ...." *Comparative Political Studies* 43:1499—1526.

Gerring, J. 2011. *Social Science Methodology: A Criterial Framework*, 2nd edition. Cambridge: Cambridge University Press.

Gerring, J., and J. Seawright. 2008. "Case-selection Techniques

in Case Study Research: a Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61:294—308.

Glynn, A.N., and Quinn, K.M. 2011. "Why Process Matters for Causal Inference." *Political Analysis* 19:273—286.

Green, D.P., S. E.Ha, and J.G. Bullock. 2010. "Enough already about 'Black Box' Experiments: Studying Mediators is more difficult than most Suppose." *The Annals of the American Academy of Political and Social Science* 628:200—208.

Hall, P.A. 2003. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Analysis in the Social Sciences*, edited by J. Mahoney and D. Rueschemeyer. Cambridge: Cambridge University Press.

Hedström, P., and R. Swedberg, eds. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.

Hedström, P., and P. Ylikoski. 2010. "Causal Mechanisms in the Socila Sciences." *Annual Review of Sociology* 36:49—67.

Holland, P.W. 1988. "Causal Inference, Path Analysis, and Recursive Structural Equation Models." *Sociological Methodology* 18: 449—484.

Imai, K., L.Keele, D.Tingley, and T.Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105:765—789.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Kreuzer, M. 2010. "Historical Knowledge and Quantitative Analysis: the Case of the Origins of Proportional Representation." *American Political Science Review* 104:369—392.

Lange, M. 2009. *Lineages of Despotism and Development: British Colonialism and State Power*. Chicago: University of Chicago Press.

Lieberman, E.S. 2003. *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. Cambridge: Cambridge University Press.

Lieberman, E. S. 2005. "Nested Analysis as a Mixed-method Strategy for Comparative Research." *American Political Science Review* 99:435—452.

MacKinnon, D. P. 2008. *Introduction to Statistical Mediation Analysis*. New York: Taylor and Francis.

Mahoney, J. 2001. "Beyond Correlational Analysis: Recent Innovations in Theory and Method." *Sociological Forum* 16:575—593.

Mahoney, J. Forthcoming. "The Logic of Process Tracing in the Social Sciences." *Sociological Methods and Research*.

Mayntz, R. 2004. "Mechanisms in the Analysis of Social Macrophenomena." *Philosophy of the Social Sciences* 34:237—254.

McAdam, D., S. Tarrow, and C.Tilly. 2001. *Dynamics of Contention*. Cambridge: Cambridge University Press.

Morton, R., and K. Williams. 2010. *From Nature to the Lab: Experimental Political Science and the Study of Causality*. New York: Cambridge University Press.

Norkus, Z. 2004. "Mechanisms as Miracle Makers? The Rise and Inconsistencies of the Mechanistic Approach in Social Science and History." *History and Theory* 44:348—372.

Pevehouse, J. 2005. *Democracy From Above: Regional Organizations and Democratization*. Cambridge: Cambridge University Press.

Porpora, D.V. 2008. "Sociology's Causal Confusion." In *Revitalizing Causality: Realism About Causality in Philosophy and Social*

Science, edited by R.Groff. London: Routledge.

Roberts, C. 1996. *The Logic of Historical Explanation*. University Park: Pennsylvania State University Press.

Robinson, J.A. 2006. "Economic Development and Democracy." *Annual Review of Political Science* 9:503—527.

Rohlfing, I. 2008. "What You See and What You Get: Pitfalls and Principles of Nested Analysis in Comparative Research." *Comparative Political Studies* 41:1492—1514.

Rueschemeyer, D., J. D. Stephens, and E. H. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.

Sambanis, N. 2004. "Using Case Studies to Expand Economic Models of Civil War." *Perspectives on Politics* 2:259—279.

Schneider, C.Q., and I. Rohlfing. 2011. "Combining QCA and Process Tracing in Set Theoretic Multi-method Research." Manuscript.

Sekhon, J. 2004. "Quality Meets Quantity: Case Studies, Conditional Probability, and Counterfactuals." *Perspectives on Politics* 2: 281—293.

Sekhon, J. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." *Annual Review of Political Science* 12: 487—508.

Slater, D., and B. Smith. 2010. "Economic Origins of Democratic Breakdown? Contrary Evidence from Southeast Asia and Beyond." Manuscript.

Snow, J. 1855/1965. *On the Mode of Communicaiton of Cholera*, 2nd edition. London: John Churchill.

Stinchcombe, A. L. 1991. "The Conditions of Fruitfulness of Theorizing about Mechanisms in Social Science." *Philosophy of the Social Sciences* 21:367—388.

Van Evera, S. 1997. *Guide to Methods for Students of Political Science*. Ithaca: Cornell University Press.

Waldner, D. 2011. "Process Tracing and Causal Mechanisms." In *The Oxford Handbook of the Philosophy of Social Science*, edited by H.Kincaid. Oxford: Oxford University Press.

9 反事实

我们准备支持这样一项反事实的宣称，即在已知 20 世纪 80 年代权力实质分布的情况下，快速没落中的苏联即便受老思想家的领导，也会在冷战时期最渴求和平。

——兰德尔·施韦勒(Randall Schweller)
威廉·沃尔福思(William Wohlforth)

几乎不可想象的是，对于戈尔巴乔夫的总书记位置的任何一位竞争者而言，他们哪一位能着手，更不用说贯彻他那大胆的国内外改革。

——罗伯特·英格利希(Robert English)

9.1 导言

反事实是社会科学方法论中的许多不同议题的核心。本章关注的是，当在定性与定量传承中进行因果推断时，反事实扮演着什么角色。在前一章中，我们讨论了如何用反事实来界定因果关系本身(参见第 6 章)。①

① 在定量传承中有一种“反事实路数”被称为 Neyman-Rubin-Holland 模型，因为它开始于一个个体个案 i 的反事实。在哲学中，反事实的因果关系定义有着一个悠久而著名的历史(例如，Lewis，1973；Collins et al.，2004)。在定性方法论中，必要条件和反事实被看做有内在联系的(例如，Goert and Starr，2003)。

延续两种传承的主题，我们认为，反事实分析在定性传统中是一种重要的因果推断模式，但是它在定量传统内部并不常使用。这种差异的原因在于如下事实，即一个反事实的陈述内含着"与实际发生相反"这样的陈述。这些通常是关于单个个案中的一个特殊原因和结果的陈述，不涉及一个总体平均而言会如何发生。例如，本章的引语就针对一个特殊事件提供了经典的反事实论证。在这个个案中，该反事实宣称的是，如果像戈尔巴乔夫这样的新思想家没有掌权，将会发生什么，而在讨论观念抑或实质权力分配在解释冷战结束的原因方面扮演着什么角色时，上述反事实的宣称对于整个讨论至关重要。

在针对一个特殊个案进行反事实的宣称时，人们通常需要针对该个案进行个案内分析。由于定性学者对解释个体个案感兴趣，因此，为了评价反事实宣称，自然会进行个案内分析。他们在不要求彻底重写历史的情况下，设想某个原因未曾发生过(或者以不同的方式发生)(Weber, 1949; Fearon, 1991)，在此意义上，他们筹划的反事实是"可以想象的"。反事实分析本身的执行显然要考虑到这样一个"可能的世界"(Lewis, 1973)，即其中缺少前因或有不同的前因。

相比之下，鉴于定量研究者对任何特定的个案不感兴趣，因而他们也不常针对特殊个案进行反事实分析。当运用反事实时，他们的目的主要是阐释一个总体的因果模型。例如，定量学者有时使用反事实来说明有关 X 对 Y 的效应的总体平均量(average magnitude)情况。这种传统下的反事实通常涉及对一个个案中的异常变化进行假设——例如，如果一个个案的 X 值从非常低变到非常高，那么估计该个案会发生什么？(相关的例子和批判，参见 King and Zeng, 2006, 2007)。

9.2 因果推断的基本问题：不同的回答

除了一些例外(如 Pearl, 2000)，目前统计学关于因果推断的观点都始于一个个案的反事实。由于不可能实际观测到个体个案的反事

实，因此，学者面对着因果推断的基本问题：

> 因果推断的基本问题。不可能观测到同一个单元的 $Y_t(i)$ 与 $Y_c(i)$ 的值，因此，不可能观测到 t 对 i 的影响。（Holland，1986：947）

因果推断的基本问题是一个反事实的问题。对于任何给定的个案来说，不可能让历史重演而使该个案在 X 上有不同的值。

定性与定量传统的重要差异在于它们对于这个问题的回答。定量学者通过聚焦于平均效应的大样本的解来解决这个问题。因此，尽管反事实在定量传统中恰好是因果性定义的核心，但反事实分析却不作为选项来进行因果推断。关于因果性的 Neyman-Rubin-Holland 反事实观点并没有将自身引向对个案的明确的反事实分析。

一旦利用统计方法计算出 t 的平均效应，研究者便可以使用这个参数估计值对总体中的任何一个特定的单元进行反事实分析。不过，本章的要点是认为，个体反事实在因果推断（即估算平均因果效应）过程中不起任何作用。人们不借助反事实分析也能获得估计的平均因果效应。统计结果在先，而非在任何反事实分析之后。

相比之下，对因果推断基本问题的定性回应是用普遍知识和个案内分析来进行反事实地分析：如果假定 X 在一个特殊个案上取了不同的值，将发生什么。定性分析者认为，某些特定的要素是否为一些特定个案结果的原因，有关这方面的假设是不能用已有的普遍知识和关于一个特殊个案的细节知识来检验的。反事实分析是产生关于特定个案的因果推断的尝试的一部分。

简单来说，定性和定量学者都认为，对个体个案进行反事实分析是可能的。然而，在定量范式中，这种分析并不是进行因果推断的部分。相反，任何反事实分析遵循的都是估计平均因果效应，用它来讨论已经做出的因果推断。相比之下，定性学者关注的是个体个案，并用现存的普遍知识和个案内分析对该个案的反事实进行推理。在该范式中，反事实分析常常是因果推断进程本身的一部分。

9.3 建构反事实

9.3.1 统计步骤与极端反事实

为了展示统计学的反事实路数，一种有益的方式是考察在 logit 或概率单位分析中，在评价一个既定的参数估计量时有什么标准的做法。在这些模型中，通常没有直接显明的方式对因果效应的规模进行概念化。分析者的一种解决方案是使用反事实来提供一种审视，即考量 X 对 Y 的效应量级。其标准步骤大致如下：

1. 将全部控制变量（即除了所关注的反事实变量 X 以外的所有变量）都设定为均值或中位值——均值最可能被选。对于二分变量，则使用众数。

2. 将反事实的先导变量 X 设定为最小值；一种比较保守的做法是取 25% 的四分位数或取值可能为低于均值的 1 到 2 个标准差。

3. 令 X 值从最小值变到最大值（一种比较保守的做法是用 75% 的四分位数或者高于均值 1 到 2 个标准差），这就是反事实。

4. 使用统计模型和参数估计法，根据 Y 的概率变化来估计反事实。

5. 用反事实中 Y 的概率变化来解释 X 对 Y 的因果效应量。

无数文章与研究者都使用了这个步骤，虽然分析者通常没有将这种做法与反事实推理明确地联系在一起。

在这些步骤中，X 从极低值变到极高值。因此，人们可以说，分析者遵循的是一种“最大重写实践”（maximum rewrite practice）或者“极端反事实路数”（extreme counterfactual approach）：反事实涉及 X 的最大或极端变化。这种实践可使研究者醒目地展示 X 的某种变化对

Y 的潜在影响。

当定量研究者从事这种反事实推理时，他们通常不将反事实与任何特定的历史个案联系起来。同样地，对于正在关注的 X 来讲，他们很少追问 X 的一种既定改变有怎样的历史可能性（historical plausibility）。如费伦注意到，这种改变是否现实并不要紧，因为它仅用于模型分析的目的：

> 在回归分析和其他检验因果假设的统计方法中，有人推断，如果样本中有任意一个特定的个案在一个自变量上取了不同的值，因变量便可能分化为在跨个案之间相同的一个系统成分（systematic component）和一个随机成分（random component）。至于能否假定任何特殊的个案在自变量上取不同的值，人们甚至永远不能深思这种假定是否真实，在历史上是否可能。（Fearon，1996：61）

可见，在定量传统中，针对个体个案的反事实是否可能，这个问题仍未引起注意。金和曾（King and Zeng，2006，2007）有力地论证到，可能的反事实需要数据来界定。数据内部的反事实是可信的；数据之外的推断则存疑。数据必须提供现实世界中的例子，用它们进行反事实分析。换言之，一个个案在某个自变量上的反事实变化对于该个案的历史专家来说可能不会发生，但是，只要数据提供了良好的比较个案——基本上是匹配的个案——可用来估计该变化产生的效应，那么这也不会太令人困扰。事实上，在定量传承中，人们可以说，任何给定的反事实都取决于该反事实距离观测数据有多远。

9.3.2 定性步骤与最小重写规则

在定性传承中，典型的反事实是关于一个特定——通常在历史上重要的——个案的一种声称。有人问，如果原因 X 在一个实际感兴趣的个案中取了不同的值，那么某个给定的结果 Y 会怎样变化。研究者

会想象这个原因事件不发生(或以不同的方式发生),并探索结果是否仍然会发生(或以同样的方式发生)。[1]在反事实假设中,通常将结果设定为一个结果的缺失或否,而非关于一个特定的正向结果的精准预测(参见 Fearon, 1996)。在许多例子中,原因 X 是个小事件,但它被认为具有重要的后果,虽然不完全如此。

定性研究与特定的反事实个案的分析有着天然的联系。如果有人提出,X 是一个特殊的个案中 Y 的成因,人们自然会询问,如果 X 有所不同,将会发生什么。事实上,如果认为 X 是 Y 的一个必要条件,该假设要想成立,那么 X 的反事实缺失必然带来 Y 的缺失。同样,在个案层次中,INUS 的缺失通常会导致结果的缺失。[2]

不出意外,在进行反事实分析时,如何使因果推断的影响力最大化,定性方法论者在这方面引领着相关的讨论。至早可追溯到韦伯(1949),学者提出了"最小重写"原则("minimum rewrite" rule),认为反事实应该要求尽可能少改变已知的历史记录(Stalnaker, 1968:104; Elster, 1978; Tetlock and Belkin, 1996b:23—25; Reiss, 2009)。在已知某个案的特征的情况下,反事实的先行条件必须可理解并且看上去可信。要基于关于特定个案和宽泛理论方面的知识和来自其他研究领域的普遍发现来评价可信度。应该避免"奇迹"式的反事实先行条件。通常来讲,最可能的反事实先行条件涉及"小"改变(例如,假设一位领导人未被暗杀),这也是为什么定性研究者在进行反事实分析时常常考察小事件的原因(Fearon, 1991:193)。类似的是,反事实的结果通常必须在一个"可能的世界"中发生,该世界不会完全不同于现实世界(参见 Lewis, 1973)。如费伦所说,"当我们试着解释为何某个事件 B 发生的时候,会潜在地设想有一个对照空间,在这个空间里没有 B,其余部分则与 B 存在的世界相似"(Fearon, 1996:57;参见 Garfinkel, 1990)。

---

① 有时候,研究者会考察一个否事件(例如,不发生战争),此时所谓的反事实就指的是一个正向事件(如战争)。

② 有一个例外是这样的情况,即其结果是由多重充分条件构型(multiple sufficient condition configurations)的存在而由多因素决定的。

好的反事实能够引导定性研究者,帮助他们指定证据,以便在解释特定的结果时与过程追踪一起使用(参见第8章)。例如,人们可能反事实地提出假设,认为如果戈尔(而非布什)赢得了2000年的总统大选,美国将不会发动伊拉克战争(参见 Harvey, 2012)。从历史角度讲,这个反事实貌似可信,它提请我们注意到个体领导人在推动外交政策上的影响。它引导研究者进一步考察布什和戈尔的差异,并区分出布什有哪些独特的信念和抉择最终导致战争决策。这也激励分析者考虑到备择的论证,例如系统性的压力会使得即使戈尔当选总统也会命令美国攻击伊拉克。相比之下,下面这个假设就不是非常有用了:如果特蕾莎修女(而非布什)赢得了2000年总统大选,美国就不会发动伊拉克战争。如果世界没有彻底的改观,这种事件的发生就是不可想象的,并且这种分析提供的指针太少了,因而不能区分出到底布什总统的领导力的哪个方面对于战争的发动来说是关键。

如此看来,我们找到了定性与定量研究中提出的典型反事实的根本差异。前一种传统提出的反事实遵循着最小重写原则,并且从历史角度讲对于正待分析的个案来说是可能的。这个标准直接遵循着这样一个事实:定性研究者试图用反事实分析对一个特定的历史个案进行因果推断。相比之下,定量传统中的学者则提议那种涉及最大重写原则的极端反事实。这样做的原因在于他们将反事实用于模型分析和解释目的,并非因果推断。

9.3.3 什么是合理的反事实？个案间与个案内的回答

金和曾(2007)在一篇重要文献中指出,定量研究中可能有一个反事实,即数据集合中有一些个案与所提议的反事实相似。他们(像一些哲学家那样)不讨论最接近的"可能的世界"(possible worlds),而关注在数据中发现的最接近的真实世界(actual worlds)(参见 Mikkelson, 1996)。当数据不支持这些反事实时,他们便批判。其中的一个例子涉及加拿大的国家失灵(因变量):

> 我们的第一个极端反事实是，假定加拿大在 1996 年成为一个威权政体，但是它的其他变量值仍然是实际值。如所期望，我们发现该极端反事实处在观测到的数据（在控制变量上的）凸包（convex hull）以外，因此需要外推。换言之，我们固然可以追问如果 1996 年加拿大成为威权政治会发生什么，但我们不能用历史作为引导，因为与加拿大在其他可测的特质方面相似的威权政体的实例在世界上是不存在的。（King and Zeng，2007：192）

金和曾用跨个案的术语界定了一个可能的反事实。如果有其他个案与反事实相似，该反事实就是合理的。

相比之下，定性研究中运用的最小重写原则是一种个案内（within-case）的观念：对于个体个案来说，反事实提议是否可信？例如，加拿大在 1996 年推行威权政治，这个反事实就不合理，因为它违反了最小重写原则。要彻底改变加拿大的历史才能得到这个反事实的先导条件。当然，人们可以用一些跨个案的证据来评估个案内的反事实，但是对其合理程度的判断最终还是一个个案内的决定，它在很大程度上依赖于分析者拥有的关于该个案及其历史的知识。

为了进一步探究这种观点，我们可以考察有多大可能性去发现这样的个案，即从完全威权变成了完全民主。例如，对于政体数据*来说，一个国家在相对较短时期内有多大可能从－10（威权主义）变成 10（民主）？统计的回答是截面式的：有没有哪些国家的政体值（Polity value）是 10 及－10，但在其他变量上相类似？①定性研究则涉及个案内分析：这种变化对于一个特殊个案来说有多大的可能性？定性研究者不完全依赖于类似的实际个案能否被发现（虽然这条信息通常有一些

* 政体数据系列（The Polity data series）是政治科学中广泛使用的数据库，其最新版本是政体数据第四版，它包括了所有人口数大于 50 万的国家 1800—2013 年间每一年的民主程度信息编码值。针对每一年的每个国家计算其政体得分（Polity Score）：－10 到－6 对应的是专制政治，－5 到 5 对应半民主，6 到 10 对应民主。其网站参见 http://www.systemicpeace.org/polity/polity4.htm。——译者注

① 事实上，这类最大改写并不被数据支持。政体数据表明，国家有可能在一段相对较短的时间内从 10 分增加到 15 分，但是不可能完全到达 20 分。

作用）。某个基于截面数据上的反事实可能注定是不可信的，但一旦定性研究者针对一个特定个案来分析时，它却可能是可信的。定性研究者也可能基于个案内知识认为一个反事实是不可信的，即使该反事实不能通过截面式的统计数据所排除。[①]

9.4 结论

在定性研究中，反事实分析居于个案内因果分析的核心。对于个案研究来说，反事实分析是研究者在评价假设时与过程追踪法共同使用的一个主要工具。他们在反事实的假定下重演一个个案的历史，以便决定一个既定的因素是否发挥了它假设具有的因果作用。这些反事实实验的结果能够强烈地影响到由个案间分析所产生的结论。

虽然反事实在统计学传承中指引着因果性的定义（参见第 6 章），处于该传承中的研究者通常还是不能着手历史个案的反事实分析。他们选择的是总体中（潜在地假设的）代表性个案，而不是特别关注的个体实际观测值。他们不用理论和已有的发现来重现历史，因此并不判断如果某个个案在一个原因变量上取了不同的值，将发生什么后果。相反，在这种传承中，表达反事实的目的是解释统计估计的结果，做出有关因果效应的普遍论断。

参考文献与推荐读物

毋庸置疑，在社会科学研究中，有关反事实的研究最早见于韦伯（Weber，1949）。近期关于反事实分析的重要文献包括费伦（Fearon，1991，1996）；泰洛克和贝尔金（Tetlock and Belkin，1996a）。有关社会

① 定性研究者也许相信，对于一个个案来说，方向上的改变（例如从完全威权到完全民主）是可信的，而反向的改变却不现实。这个观点值得考虑，因为它体现了定性研究中反事实与因果不对称观念之间的联系（参见第 5 章）。

科学中的反事实分析的定性例子，参见泰洛克和贝尔金(Tetlock and Belkin，1996b)；格尔兹和利维(Goertz and Levy，2007)以及勒博(Lebow，2010)。历史学家关于反事实的研究，参见费希尔(Fischer，1970：15—21)，麦克莱兰(McClelland，1975)和弗格森(Ferguson，1999)。哲学上的研究可以在刘易斯(Lewis，1973)，萨蒙(Salmon，1994)和伍德沃德(Woodward，2003)中找到。法学中的反事实研究，可以参见哈特和奥诺雷(Hart and Honoré，1985)的经典著作。

将反事实分析与跨个案分析结合起来的研究并不多。从集合论角度的重要研究有拉金(Ragin，2008：第8—9章)。从统计分析的角度看，至于哪些类型的反事实能够和不能被评估，金和曾(King and Zeng，2006，2007)，戴维(Dawid，2000)提供了良好的讨论。也可参见珀尔(Pearl，2000)。

Collins，J.，N. Hall，and L.A. Paul，eds. 2004. *Causation and Counterfactuals*. Cambridge：MIT Press.

Dawid，P. 2000. "Causal Inference without Counterfactuals(with Discussion)." *Journal of the American Statistical Association* 95：407—450.

Elster，J. 1978. *Logic and Society*：*Contradictions and Possible Worlds*. Chichester：Wiley.

Fearon，J. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43：169—195.

Fearon，J. 1996. "Causes and Counterfactuals in Social Science：Exploring an Analogy between Cellular Automata and Historical Processes." In *Counterfactual Analysis in World Politics*，edited by P.Tetlock and A.Belkin. Princeton：Princeton University Press.

Ferguson，N.，ed. 1999. *Virtual History*：*Alternatives and Counterfactuals*. New York：Basic Books.

Fischer，D.H. 1970. *Historians' Fallacies*：*Toward a Logic of Historical Thought*. New York：Harper & Row.

Garfinkel，A. 1990. *Forms of Explanation*：*Rethinking the*

Questions in Social Theory. New Haven: Yale University Press.

Goertz, G., and J.Levy, eds. 2007. *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*. New York: Routledge.

Goertz, G., and H.Starr, eds. 2003. *Necessary Conditions: Theory, Methodology, and Applications*. New York: Rowman & Littlefield.

Hart, H.L.A., and T.Honoré. 1985 .*Causation in the Law*, 2nd edition. Oxford: Oxford University Press.

Harvey, F. 2012. *Explaining the Iraq War: Counterfactual Theory, Logic and Evidence*. Cambridge: Cambridge University Press.

Holland, P.W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945—960.

Khong, Y.F. 1996. "Confronting Hitler and its Consequences." In *Counterfactual Analysis in World Politics: Logical, Methodological, and Psychological Perspectives*, edited by P.E.Tetlock and A.Belkin. Princeton: Princeton University Press.

King, G., and L.Zeng. 2006. "The Dangers of Extreme Counterfactuals." *Political Analysis* 14:131—159.

King, G., and L.Zeng. 2007. "When can History be our Guide? The Pitfalls of Counterfactual Inference." *International Studies Quarterly* 51:183—210.

Lebow, R.N. 2010. *Forbidden Fruit: Counterfactuals and International Relations*. Princeton: Princeton University Press.

Lewis, D. 1973. *Counterfactuals*. Cambridge: Harvard University Press.

McClelland, P.D. 1975. *Causal Explanation and Model Building in History, Economics, and the New Economic History*. Ithaca: Cornell University Press.

Mikkelson, G. 1996. "Stretched Lines, Averted Leaps, and Excluded Competition: a Theory of Scientific Counterfactuals." *Political Studies* 63:S194—S201.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Reiss, J. 2009. "Counterfactuals, thought Experiments, and Singular Causal Analysis in History." *Philosophy of Science* 76: 712—723.

Rubin, D.B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:688—701.

Rubin, D.B. 1990. "Comment: Neyman(1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5:472—480.

Salmon, W.C. 1994. "Causality without Counterfactuals." *Philosophy of Science* 61:297—312.

Stalnaker, R.C. 1968. "A theory of Conditionals." In *Studies in Logical Theory: Essays*, edited by N.Rescher. Oxford: Blackwell.

Tetlock, P. E., and A. Belkin, eds. 1996a. *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*. Princeton: Princeton University Press.

Tetlock, P. E., and A. Belkin. 1996b. "Counterfactual thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives." In *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*, edited by P.E.Tetlock and A.Belkin. Princeton: Princeton University Press.

Weber, M. 1949. "Objective Possibility and Adequate Causation in Historical Explanation." In *The Methodology of the Social Sciences*, edited by M. Weber. New York: Free Press.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

第三部分　概念与测量

10 概念：定义、指标与误差

事物的本质在于……没有它，该事物便既不成立，也不能认为它成立。

——J.S. 穆勒(J.S. Mill)

总之，可以尝试用一个既定的指标对某个系统化的概念进行操作化，当从该指标中引出的一些值可以依据此系统化概念进行有意义地解释的时候，测量才是有效的。

——罗伯特·阿德科克(Robert Adcock)
戴维·科利尔(David Collier)

10.1 导言

从事着两种传承研究的学者们采用不同的路数来处理概念化和测量问题。其中一些路数之间的差异不是特别令人惊讶。定性学者早已就概念的意义进行了长期、集中、"冗长"的讨论。就此而言，他们像(政治)哲学家，也在概念分析上花费大量时间。相比之下，定量学者的统计模型则需要数据。因此，他们集中关注定量测量的质和量。他们在概念上耗时较少，在操作化、测量和得到的数据集上费时较多。

本章关注有关概念化和测量的定量和定性路数之间的两个重要区别。第一个区别涉及的问题是，将概念界定或概念测量中的哪一个赋予相对重要的地位。定性研究者最关注定义问题及其概念的意义。他

们通常接受语义学的路数，努力确定哪些界定性的属性构成了一个概念的固有本质。相比之下，定量学者的关注点主要放在潜变量的测量上。他们寻找与所研究的潜变量有关系的优良指标。

第二个大的区别涉及误差和个案编码。当定性研究者处理那些拥有极值（如那些接近理想型）的个案时，会感觉到最有把握。他们对那些其取值在全域（range）中间的个案最无把握。相比之下，定量研究者对有极值的个案最没有把握，对接近均值的个案最有把握。在这两种传承中，将定义和指标中的哪一项赋予相对重要的地位，这一点与上述差异是相关的。定性研究者将数据与定义匹配起来，因而，对于那些其定义显然与数据匹配或不匹配的实例来说，定性研究者最有自信。相比之下，定量学者使用的是指标，对于那些其取值远离作为一个整体的总体中典型值的个案来说，他们最不自信。

10.2 对特征及指标的界定

在定性传承中，关于概念的讨论和争论涉及语义学，即涉及概念的意义。在追问并讨论概念的定义是什么的时候，一个完全标准的问法就是追问其意义。例如，有人可能会问，“你对福利国家的定义是什么？”一个典型的定性回答是给出一张构成该概念的属性或特征的列表。这个做法没什么太神秘之处，因为它基本上就是词典所为之事。词典和定性传承都试图设定一个词条的本质特征。现有来自哲学的一个常用例子，关于“铜”的一个好的定义会给出其化学特征，这些特征描述了铜的性质，包括任意一些因果力（causal powers）。

在定量传承中，关于概念的讨论和争论集中在数据和测量问题上，很少关注语义学和意义。虽然关于概念界定的讨论对于收集此概念的数据来说通常是必要的，但是它并非关注的焦点。在有关定量测量的文章中，概念部分甚至根本不存在。相反，研究者关注的是概念的操作化和测量。操作化通常包括发现一些由数值数据构成的指标，这些指

标要与不可测量的潜变量有关。一旦找到这样的指标，就可以利用简单或复杂的测量程序对个案进行编码。这些测量程序的范围从简单的相加到复杂的贝叶斯式潜变量模型。聚合程序或测量模型会生成与所研究的概念（变量）相关的个案得分。

作为此类定量方法的一个例子，可以考虑 GTD 恐怖主义数据库（CETIS，2007）*，该数据库目前经常用在恐怖主义的统计研究中。众所周知，恐怖主义是一个有疑问的概念（参见 Schmid and Jongman，1988 给出的几十种定义）。阅读 GTD 编码本可见，导言部分清楚地承认此概念的不确定性，但是几乎所有的编码本都与数据相关。一旦承认此概念难以定义，定义问题就不在考虑之列了。在不参考定义问题的情况下，要继续讨论数据。事实上，要想区分恐怖主义的实际的 GTD 定义，必须阅读附录。

定性学者并不需要大量的个案数值数据，他们自由地讨论概念及其界定性的属性。这种自由的一个危险在于增加了概念的复杂性。定性的定义可能很长、难懂，甚至有拜占庭式的特征（Byzantine in character）。我们青睐的一个例子是施米特（Schmitter）给出的关于法团主义（corporatism）的有影响的定义：

> 可以将法团主义定义为一个利益代表体系，在这种体系中，其构成单位（constituent units）被纳入少数几个单一性的、强制性的、非竞争性的、有层级秩序以及按功能区分的范畴类别之中，它们得到国家的承认或许可（如果不是由国家创造的话），并在各自的范畴类别中享有垄断的代表权，作为交换，它们在领导选择以及需求和支持的表达方面接受国家的某些控制。（Schmitter，1974：93—94）

这个定义有很多不同的属性，其中一些包含在其他属性之中。如果有人试图将施米特的定义分解成个体的特征，其中可能存在 10 个或

* GTD 是全球恐怖主义研究数据库（Global Terrorism Database）的缩写，它是一个开放源代码的数据库网站，这里记录了从 1970 年到 2013 年世界各地的恐怖事件信息，并且不断更新各种恐怖事件，其网址参见 http://www.start.umd.edu/gtd/。——译者注

更多的特征。此外，不同的人可能给出不同的列表。

在定量传承中，对有关法团主义数据的编码过程涉及指标的使用。这些指标可能在概念的定义中没有明确地提及。例如，劳工集中势(labor centralization)被用作法团主义的定量指标(参见 Kenworthy, 2003 对其他定量指标的讨论)。至于该指标如何符合施米特的定义中的抽象语言，则是一个需要解释的问题。一般来说，从概念到具体数据的变化将几乎总是涉及显著的简化。通常这种简化隐含着需要重新定义一个概念，使之包括更有限的定义维度。

对定性研究者而言，指标如果不能很好地代表一个概念的全部界定性的属性，这就值得关注了。对于这些研究者而言，一个概念的诸多属性都是那些真正构成此概念的必要特征。因此，每一个属性都要测量。定性研究者拒绝极端的简化模式。他们相信，对概念的界定必须独立于对数据的考虑。一个概念的定义不应该受到可测量此概念的数据的驱使。

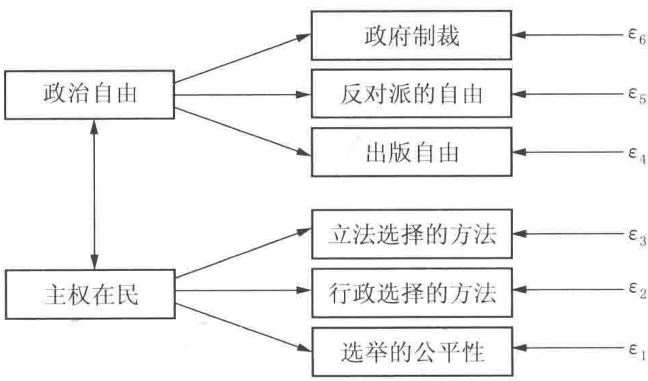
与构成一个概念的诸多属性不同的是，定量指标是可选、可替换的，未必是定义性的。不同的指标都是同一个概念项的测度。特赖尔和杰克曼(Treier and Jackman, 2008)对民主的政体测度(Polity measure)很好地展示了界定性的属性和指标之间的差异。政体数据从 5 个属性角度来定义民主，它认为每一个属性都是民主的内在特征。然而，对于特赖尔和杰克曼来说，这些属性无非是潜在的民主概念的指标而已。在这 5 个指标中，有 2 个不符合他们的方法论的统计要求，因此被丢弃。因此，他们最终的民主测度只用了 5 个政体维度中的 3 个。

因此，定性研究者所给出的界定性特征对于定量研究者来说可能是一个指标。另举一例，博伦和格朗让(Bollen and Grandjeans, 1981)使用选举的公平性作为政治民主的一个指标。对定性研究者而言，公平选举经常被看做民主的一个界定性的固有属性。它不是可选的：这个属性对于民主来说是必要的属性(Bowman et al., 2005; Mainwaring et al., 2001)。

在讨论概念的定性文献中，指标的语言通常被诸如一种民主的最低要求这样的语言所替代。一个好的例子是科利尔和列维茨基关于缩

减子类(a diminished subtype)的有影响力的见解："例如，将'有限选举民主'(limited-suffrage democracy)和'监护式民主'(tutelary democracy)都理解为不完整的民主个案，因为它们缺少民主的一个或多个界定性属性。"(Collier and Levitsky, 1997:436—437)如果源概念(root concept)的诸多属性具有备选性，对于该源概念中的概念成分资格来说不是必要的，那么缩减子类思想恰恰意义不大了。

一个相关的差异涉及定性研究中的属性/概念与定量研究中的指标/变量的对比。在定性传承中，属性和概念之间有语义学的、概念上的关系。在定量传承中，指标和潜变量之间是一种因果关系：用标准眼光来看，潜变量产生了指标(Bollen, 1989)。①



资料来源：Bollen and Grandjean, 1981。

图 10.1　关于政治自由和主权在民(popular sovereignty)的潜变量模型

图 10.1 展示了一个典型的潜变量模型。图中的因果箭头从潜变量指向了指标。这有助于理解这样的想法，即指标是可替换的因素。潜变量可能是许多不同事情的起因，学者只选出其中的一些。如上所述，当研究者将这些指标看成是界定性属性的时候，就会产生跨越两种传承的张力。在定性传承中，界定性属性不能与感兴趣的概念有因果联系；它们甚至不能从时间上与概念分离。它们就是概念。

① 对因果箭头的建模也可以取另一个方向(如 Bollen and Ting, 2000)，但是这种情况很少见。

测量应该建立在因果理论的基础之上，这种思想已有长期的、有名的历史。亨普尔（Hempel，1952）用自然科学作为其关注点，给出了这种联系。例如，温度计之所以能测量温度，其有效性取决于热膨胀的因果理论。在社会科学中，同样的想法也支撑着众多关于潜变量和测量的文献（Bollen，1989）。我们的目的不是质疑这样的观点，即指标和潜变量应该因果相关。在定量传统中，对于很多现象来讲，这种见解完全可以理解。相信一个人的智力——潜变量——可能影响其在智力测验上的表现，这看起来是合理的。同样，一个人的政治意识形态可能影响其对政治问卷的回答，这看起来也说得通。

不过，从定性的立场看，关键问题仍然是解决所关注的概念的意义问题。这些研究者将通过追问"到底什么是智力（或政治意识形态）的定义"而向定量学者施压。概念可以根据用来测量概念的指标来定义，他们会对任何此类回答都不满意。①从定性的视角看，指标的性质必须总要用正在测量的概念的意义维度来评估。很多定量研究者会在原则上同意这一点，但是这种传承中的关注点将讨论引导到更集中关注测量和指标问题，而不是关注意义和定义问题。

10.3 误差

当对个案进行编码时，在对待我们的知识的性质方面，定性和定量范式展示了重要的差异。在一种传承里被认为是可进行精确描述和编码的好个案通常在另一种传承里被认为是不好的。定量研究者假定的那些有高度测量误差的一类个案，恰恰通常是定性研究者假定的有最小测量误差的个案，反之亦反。

为了探索这种差异，我们可以先澄清"误差"（它是所有统计学的中

① 操作主义教条认为，概念的含义要在用来测量它们的指标和方法中寻找。大多数哲学家和社会学家都反对这一教条。

心）和“模糊性”（fuzziness）（它在定性研究中是一种重要的想法）之间的关系。这两个概念初看起来可能很相似。当定量研究者听到“模糊的”这个词的时候，他们可能最开始相信它暗含着缺乏精确性，进而意味着“不确定性”和“误差”。但是事实上，将“误差”和“模糊性”进行类比是相当误导人的。这种对比是统计学和模糊逻辑学之间转译问题的一个实例。

在统计学中，误差估计涉及我们的知识性质。的确，统计学之所以不同于关于世界的其他数值估计方式，恰恰在于它包含着一种随机元素，从而允许对我们的估计量的精确性有所评说。相比之下，模糊集成员值（fuzzy-set membership values）是对世界特征的陈述。比如，如果某人说一个个案的模糊集成员值是 0.75，他就是在宣称该个案具备的一个经验性质——即此个案主要但不完全隶属于一个给定的概念集合。关于误差和知识性质，人们不做任何假定或陈述。这样的宣称在任何意义上都不是概率性的。

一个更好的类比是，个案在模糊集合中拥有其得分，它可比拟给定的一次观测所对应的变量在一个定量数据集合中的值。尽管模糊集合的隶属值和变量值不能机械地相互转译（参见第 12 章），但是二者是对应的。一个大的区别是，模糊集合的隶属值通常不存在与之相关的误差估计值或不确定估计值。在模糊逻辑数学中，的确存在着“二级”模糊性（“second order” fuzziness）观点：我们的模糊隶属值又在多大程度上是模糊的（如 Klir and Yuan，1997；Arfi，2010）？这种二级模糊性从性质上讲可以很好地与误差估计值类比。然而，虽然在模糊逻辑文献中存在着这样一种数字选项，但是在定性传承中它可不是一件自然之事，它实际上很少应用在社会科学实践中。

尽管定性研究者没有给出明确的误差估计量，但他们的确时常讨论在对特殊个案进行精确地编码时遇到的困难。困难可能包括详尽地讨论针对个案的某些编码背后的推理。他们可能将决策建立在已有的文献或他们自己的专业知识基础之上。从定量的视角看，这些研究者会非常担忧少数个体观测值的特殊编码，这可能令人奇怪。在定量传承中，将时间和资源放在少量有问题的观测项通常不是一个好的投资。

在决定这些定性编码的时刻，通常要评定一个个案在多大程度上符合一种“理想型”或一个给定概念的纯粹的完备实例。这个理想型可以作为一种标准，依据它评估所有的经验个案，学者根据这个标准来“校对”(Ragin, 2008)他们的个案编码。从模糊集合的成员得分量表方面讲，理想类型居于量表的一极：理想型个案的成员得分为 1。这些个案明显拥有所讨论的概念的全部界定性特征。

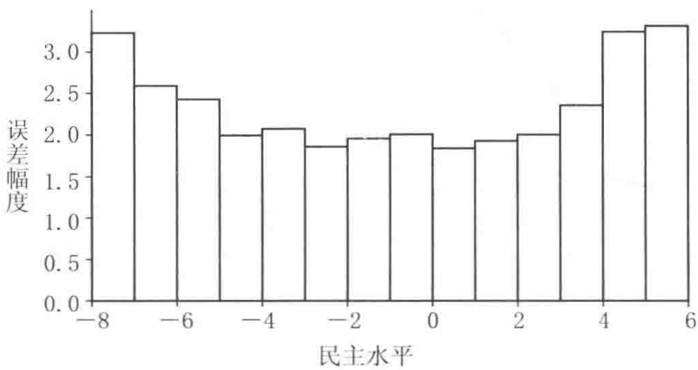
在这种定性路数中有一个总的直觉是，个案越接近理想型，越容易编码，因此，与这些编码相关的误差也小一些（Eliason and Styker, 2009; Ragin, 2008）。同样，得分为 0 的个案通常也易于编码，因为某个因素在什么时候根本就不为一个概念所包含，这通常是很清楚的。相比之下，模糊集合得分为 0.5 的个案展示出最大的模糊性，特别难以编码。因此，从模糊集得分为 1 的理想型转到最大模糊度 0.5，精确地编码的难度增加，误差更可能出现。从得分为 0.5 再移到极值 0，编码就又变得容易了，误差普遍减少。因此，在实践中，在模糊水平和误差水平之间通常大致有曲线关系。这种关系因情况而定，它取决并依赖于正在被测量的特定现象，但我们认为它可能至少在社会科学情境中是相当普遍的。

在统计学传统中，变量值和误差之间的关系遵循反向模式。对于一个连续编码的变量来说，如果询问最大的误差估计将发生在哪里，统计学给出的答案是发生在有极值(即最高或最低)的个案中。误差在一个变量的上界和下界处最大，在中间最小。学生早在接受统计学训练、考察围绕回归线的置信带(confidence band)的时候，就了解这种关系了。估计的误差在 X 的均值及 Y 的均值处最小。距离中间点越远，误差越大。

关于统计学传统中变量值和误差之间的关系，一个好的例证是特赖尔和杰克曼（Treier and Jackman, 2008）的民主测量模型。正常情况下，在一个定量数据库——如政体数据或自由之家数据（Freedom House）*——中，对民主的编码不包括任何明确的误差估计。例如，一

* “自由之家”(Freedom House)是由美国政府资助的一个非政府组织，建立于 1941 年，它支持关于民主、政治自由和人权的研究，并发布这些方面的研究报告。——译者注

个个案可能有一个民主得分为 3，另一个得分为 6，但是与那些值有关的误差项没有被估计。特赖尔和杰克曼的模型为这些估计提供了一个基础。图 10.2 呈现了特赖尔和杰克曼针对政体数据给出的民主测度误差估计值。我们从中看到了经典的置信带形状：中间最窄，两端最宽。



资料来源：Treier and Jackman，2008：图 2。

图 10.2　统计测量中的误差展示：民主水平

从定性观点来看，针对民主的这些误差估计值看起来违反直觉。对完全民主的或者明显不民主的个案的编码通常是容易的；处在中间地带的个案难以编码。每个人都同意，瑞典是充分民主的国家，朝鲜显然是不民主的国家。但是，我们应该怎么对处在边界线上的个案（如危地马拉、委内瑞拉和洪都拉斯）进行编码呢？

在探究这个问题的时候，一种方法是询问有关民主的不同数据集之间的一致程度。可想而知，这些数据集会在易于编码的个案上趋于一致，在难以编码的个案上趋于分歧。"政体数据"和"自由之家"数据编码有重叠，在图 10.3 中，我们利用国家-年代格式（country-years）报告了这两个数据编码的方差（细节参见 Goertz，2008）。当这两个数据集完全一致的时候，它们之间的方差为 0。随着二者的分歧增加，它们之间的方差也在加大。在图中，我们清楚地看到，在威权政体和民主政体这两极（即－10 和 10）方差最小；也就是说，当政体编码和自由之家编码都是一个极端威权值（得分为－10）或完全民主值（分数为 10）的

时候，不一致性就很小。随着我们移向中间的灰色地带(0 分)，可以看到对一个给定的国家-年代编码中的变异在显著增加。事实上，随着我们从 10 分下移到 0 分，方差增长近 1 000 倍(从 0.025 到 22.6)。当我们从极端威权值(－10)上移到 0 分，也发生一个大的变化，尽管增长幅度仅有 10 倍。①

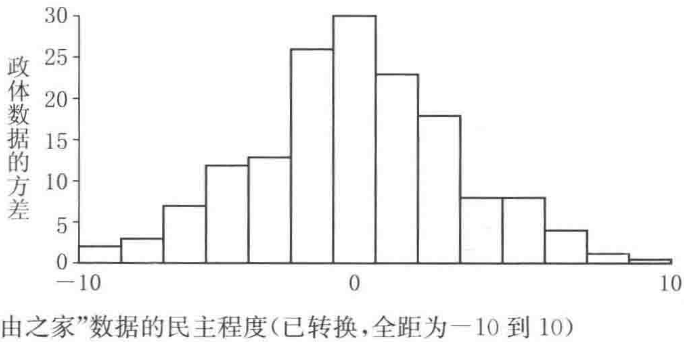


图 10.3　自由之家和民主政体量表的方差

对比图 10.2 和图 10.3，可以生动地展示两种传承及其关于测量误差位置的不同观点。图 10.2 是定量研究者通常期望发现的，图 10.3 是定性学者认为要发生的。

10.4　结论

本章讨论了推动定性和定量传统的概念和测量的不同方式。在定性传统中，总体来说，概念是通过一个语义进程建构的，其中研究者要确定构成一个概念的属性，据此来定义此概念的意义。当在这种研究中测量个案的时候，如果它就所研究的概念来说有极值，那么定性学者

① 读者可自行重评普沃斯基等人(Przeworski et al.，2000：58—59)的论证，即如果误差有着如图 10.2 所示的方差，并且民主和威权之间的分割点是 0，那么他们对民主的二分编码产生的误差就小于一个连续测度的误差。

对这样的个案最有自信。因此，当他们考察的个案明显是或明显不是所研究的概念的实例时，便发现其测量是最简单的。当个案变得像所测量的概念的模糊实例时，误差易于出现。

总起来讲，在定量传统中，要通过区分出指标来建构概念，这些指标是由所研究的概念引起的。出于各种统计学原因，误差估计值在均值上比在极值上要小；如此看来，定量学者对有极值的个案最没有把握，对接近均值的个案最有把握。

每一种传承的独特取向都有其合理性，各自都有长期的历史和背后的良好的方法论辩护。它们与这两种传统的个案取向对总体取向的本质密切相关。然而，这些根深蒂固的差异使它们之间难以（尽管不是不可能）通融，达到对概念和测量的一种综合的研究（Adcock and Collier, 2001）。这有助于解释为什么关于概念的定性文献和关于测量的定量文献并没有更多地互评。

参考文献与推荐读物

定性研究和定量研究赋予概念化、测量和最终数据以相对不同的权重，从这个角度来比较这两种传承在任何一个既定的概念或变量上的工作，这是很有趣的。几乎依据定义即可知，一篇“概念性”文章并不提供测量模型（尽管它在原则上可以提供），它通常也不对个案进行系统地编码。个体个案当然可以用做例子，但是其中并没有系统的分析。例如，关于民主“质量”的定性研究文献（如 Diamond and Morlino, 2005）关注的是概念问题，在操作化和系统的数据编码方面投入的时间较少。“数据型文章”（Data articles）[如可参见《和平研究杂志》（*Journal of Peace Researh*），其中有常规性的“数据性”文章]可能开始于讨论概念，但是与对操作化、指标、测量和作为结果的数据集的处理相比，这还是简单的。利用既有数据的测量性文章则几乎完全关注测量问题（例如，Bollen and Grandjean, 1981; Treier and Jackman, 2008）。

随着时间的推移来追踪概念、测量模型及其结果数据，这也蛮有趣。例如，尽管关于民主的政体测度（Polity measures）已经执行了几

十年时间(Gurr，1974；Jurr et al.，1990；Jaggers and Gurr，1995)，但是“政体”概念和测度并非一成不变。考察在各种不同版本的数据中间哪些国家在哪些年保持不变，这将是一件有意义之事。

Adcock，R.，and D. Collier. 2001. “Measurement Validity：a Shared Standard for Qualitative and Quantitative Research.” *American Political Science Review* 95：529—546.

Arfi，B. 2010. *Linguistic Fuzzy Logic Methods in Social Sciences*. Berlin：Springer-Verlag.

Bollen，K. 1989. *Structural Equations With Latent Variables*. New York：John Wiley & Sons.

Bollen，K.，and B.Grandjean. 1981. “The Dimension(s) of Democracy：Further Issues in the Measurement and Effects of Political Democracy.” *American Sociological Review* 46：651—659.

Bollen，K.，and K.Ting. 2000. “A Tetrad Test for Causal Indicators.” *Psychological Methods* 5：3—22.

Bowman，K.，F. Lehoucq，and J. Mahoney. 2005. “Measuring Political Democracy：Case Expertise，Data Adequacy，and Central America.” *Comparative Political Studies* 38：939—970.

CETIS. 2007. GTD2(1998—2004)：“Global Terrorism Database，Draft 1.0.” Manuscript，University of Maryland.

Collier，D.，and S. Levitsky. 1997. “Democracy with Adjectives：Conceptual Innovation in Comparative Research.” *World Politics* 49：430—451.

Diamond，L.，and L.Morlino，eds. 2005. *Assessing the Quality of Democracy*. Baltimore：Johns Hopkins University Press.

Eliason，S.，and R.Stryker. 2009. “Goodness-of-fit Tests and Descriptive Measures in Fuzzy-set Analysis.” *Sociological Methods & Research* 38：102—146.

Goertz，G. 2005. *Social Science Concepts：A User's Guide*. Princeton：Princeton University Press.

Goertz, G. 2008. "A Checklist for Constructing, Evaluating, and Using Concepts or Quantitative Measures." In *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmier, H. Brady, and D.Collier. Oxford: Oxford University Press.

Gurr, T. 1974. "Persistence and Change in Political Systems, 1800—1971." *American Political Science Review* 68:1482—1504.

Gurr, T., K. Jaggers, and W. Moore. 1990. "The Transformation of the Western State: the Growth of Democracy, Autocracy, and State Power since 1800." *Studies in Comparative International Development* 25:73—108.

Hall, P. 2003. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Analysis in the Social Sciences*, edited by J. Mahoney and D. Rueschemeyer. Cambridge: Cambridge University Press.

Hempel, C. 1952. "Fundamentals of Concept Formation in Empirical Science." In *International Encyclopedia of Unified Science*, vol.2, no.7. Chicago: University of Chicago Press.

Jaggers, K., and T. Gurr. 1995. "Tracking Democracy's Third Wave with the Polity III Data." *Journal of Peace Research* 32:469—482.

Kenworthy, L. 2003. "Quantitative Indicators of Corporatism." *International Journal of Sociology* 33:10—44.

Klir, G., and B. Yuan. 1997. *Fuzzy Set Theory: Foundations and applications*. Englewood Cliffs: Prentice-Hall.

Mainwaring, S., D. Brinks, and A. Pérez-Liñán. 2001. "Classifying Political Regimes in Latin America, 1945—1999." *Studies in Comparative International Development* 36:37—65.

Przeworski, A., M. E. Alvarez, J. A. Cheibub, and F. Limongi. 2000. *Democracy and Development: Political Institutions and Well-being in the World, 1950—1990*. Cambridge: Cambridge university Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Schmid, A., and A.Jongman. 1988. *Political Terrorism: A New Guide to Actors, Authors, Concepts, Data Bases, and Literature*. New York: North Holland.

Schmitter, P. 1974. "Still the Century of Corporatism?" *Review of Politics* 36:85—131.

Treier, S., and S. Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52:201—217.

11 意义和测量

在关于定量及其承载的标准逻辑规则的这场混辩中，我们简直忘记了：概念的构建先于定量。

——乔瓦尼·萨托里(Giovanni Sartori)

11.1 导言

不同的学术传承通常有不同的方法论词汇。一种传承中的学者可能假定能够运用他们已有的词汇来理解来自其他传承的概念，尽管如此，转译问题事实上经常会出现。在自然语言之间进行转译时，一个既定的词或观点有时候不能用另一种语言完全表达出来。同样，在社会科学中，当相同的术语被用在跨度不同的方法论传承时，它可能有不同的意义或有不同层次的重要性。一个概念在一种传承中是核心，在另一种传承中很可能仅得到少量的关注。

在本章中，我们针对与概念和测量有关的问题，探索这些转译问题是如何体现在定性和定量传承之中的。在定量传承中，人们谈及的是变量和指标。X 和 Y 通常是潜在的、未被观测的变量，需要(定量的)指标来测量。例如，可以考察"经济发展"这个变量。尽管该变量不能被直接观测，但它能用一个指标，如人均国内生产总值进行经验地测量。在实践中，定量传承中的学者可能将变量和指标融为一体。例如，他们可能交替使用经济发展和人均国内生产总值。然而，基于本章的

目的，我们应该令变量及其指标保持相当的分离。当我们说“变量”的时候，指的是一个有理论旨趣的潜在构念(latent construct)；说“指标”时，指的是测量这种潜在构念的数值数据。

定性学者缺少讨论这些问题的一种统一的方法。出于方便，或由于统计学课程和主导性词汇的存在，他们太经常使用变量-指标语言了。然而我们相信，这种语言在定性传承中会引起一个转译问题，并且不能把握研究实践。我们不从变量和指标的角度来讨论，而需要通过区分概念和数据来把握定性传承。所谓“概念”，我们意指一个范畴(或集合)，其中的个案有隶属资格，包括不同程度的资格。例如，一个标准的定性概念是“经济发达国家”。所谓“数据”，我们指的是性质和数量各异的信息，它们被用来评定个案是否或者在多大程度上成为概念的成员。一方面，在“变量”和“概念”之间有一个明显的类比，另一方面，在“指标”和“数据”之间亦然。概念和变量是词汇及相关的观念，可用它们来明确地表达理论，而数据和指标都是经验的信息，使用它们去测量概念和变量。

虽然变量-指标和概念-数据大体上看似是同一事物的两种说法，事实上它们指的是不同的关系。本章的标题“意义和测量”即总结了这些差异。对于定性学者来说，概念和数据之间的关系是一种语义学的关系，即意义。这些学者探索了如何用数据来表达一个概念的意义。相比之下，对于定量学者来说，变量和指标的关系关注的却是变量的测量。这些学者关心的是如何用指标来最佳地测量一个潜变量。

举一个例子，现在再一次考量如何用人均国内生产总值数据来分析经济发展。在定量传承中，一个标准的步骤是说，一个国家的人均GDP是其经济发展水平的一个好的指标。人均GDP数据进而变成经济发展水平的实际测度被用在统计模型中。定性研究者在分析经济发展时，可能也用人均GDP数据。然而，他们通常会询问，这些数据与其所“意味”的一个特定范畴(如经济发达国家)是如何关联的。数据在用于分析之前，通常必须经历我们所谓的“语义学转换”(参见Ragin，2008关于“校准”的论述)，如此才能更好地拟合研究者的核心概念，即经济发达国家。定量研究者在一个统计模型中分析人均GDP数据之

前，也可能对它进行转换（例如，使用经过了对数转换的数据），[①]但是这些转换将是因为统计原因被执行的（例如，由于有偏态数据），而不是为了实现与概念意义的更好匹配。

当在国家之间比较"经济发展水平"变量的最终数值测度与"经济发达国家"范畴的时候，对于特定的国家往往会产生不同的特殊值。因此，两种传承中的差异对于学者如何编码和理解特定的个案来说具有直接、具体的意含。

我们认为，模糊集分析是一种有用的工具，可用它来在形式上设定定性研究者考量概念-数据之间关系的方式。毕竟，模糊逻辑从起源上看是一种语义学的数学理论（关于它的通俗易懂的介绍，参见 Kosko，1993，McNeill and Freiberger，1994）。设计它的目的是解决与自然语言术语的建模有关的问题。经典的例子是诸如"高个子"和"富人"。对于这些概念，我们可能拥有关于个体的身高和财富的数据。模糊集分析提供了一套工具，这套工具用这种数据来说明个体在多大程度上隶属于高个子和富人范畴。

在定量传承中，如果人们想评估一个关于高个子或富人的假设，就可能在统计分析中使用关于身高或收入的数据。虽然变量和指标不需要在概念层次上融合，但是在实践中，出于统计假设检验的目的，这两者被看做一样的东西。定性学者在实践中也可能将概念和数据融合在一起，但是它遵循的是一种反常的程序。这些研究者倾向于进行一次非线性的语义转换，最终达至理解个体在多大程度上落入高个子和富人范畴内。

另一个思考这个问题的方式是，定量学者典型地假定变量和指标之间有线性关系。在指标和变量之间有一个直接的匹配。显然，任何一个指标都可能不是一个潜变量的完美测度，这就是为什么人们鼓励使用多重指标的原因。然而，也无须假定在变量和指标之间有一个非线性的关系。相比之下，在模糊逻辑内，语义学转换几乎从来不是线性的。虽然在模糊逻辑中大量地使用转换，但是，线性转换却根本不受

① 我们在第 12 章讨论一些常见的变换，如取对数。

欢迎。

在定量传承中经常使用线性，可以通过考量一对对变量看到这一点，例如发展水平对不发达水平。这两个变量通常被看做恰好互逆的：发达水平每增加一个单位，会引起不发达水平减少同样一个单位。发达和不发达是用同样的数据（如人均 GDP 数据）来测量的，这一事实强化了这种对称性。相比之下，模糊逻辑可能很好地为经济发达引发了一个语义学转换，为经济不发达引发另一个不同（如果相关）的转换。结果是，在发达国家集里一个个案的最终隶属度将与其在不发达国家集合中的隶属度负相关，但是这种关系可能不完全对称。我们不能基于一个国家的不发达分值来必然预测其在发达上的模糊集分值。对立的一对概念有非同一性（nonidentity），这个假定是标准的模糊逻辑立场，它在定性研究者中间非常普遍，例如，和平不等于不战争（参见第13 章）。

11.2 语义学变换和集合-成员函数

术语上的差异通常标志着方法论实践中的重要对比。在针对概念的模糊集合研究中，核心是“隶属函数”。为了便于展示，现考虑适用于描述人的“身高”概念。在模糊集分析中，人们询问一个既定的人在“高个子集合”中的隶属度。相比之下，在定量分析中，人们可能询问一个人在“身高水平”变量上的得分。为了对个案编码，定性和定量研究者都可能转向关于身高的数据。但是这种数据可能以不同的方式被使用。

按照惯例，模糊集合变量的范围是从 0 到 1，即[0, 1]。从严格的数学角度来说，[0, 1]的限定是随意的，因为处在$-\infty$到$+\infty$范围内的任何值都可以被重新调整到[0, 1]内。实际上，我们可以通过简单的变换将数据重新调整到[0, 1]内。在将一个连续数据集变换到[0, 1]范围内的时候，最显著的方法是减去其中的最小值，然后再除以数据集

的全距。这使得数据集中的最大值变成了 1,最小值变为 0。从统计学上讲,这是一种线性转换,并且总体上对于统计结论来说不会有差异。因此,在许多情况下,从统计学的观点看这样做是无意义的。①

就模糊集合的隶属性研究路数来说,它的一个理论和方法论特征是强行将全部数据转换到[0, 1]内。虽然变量变换在定量研究中很常见(参见第 12 章),但它们经常完全是可选的。而在模糊集分析中,它们通常不可避免。

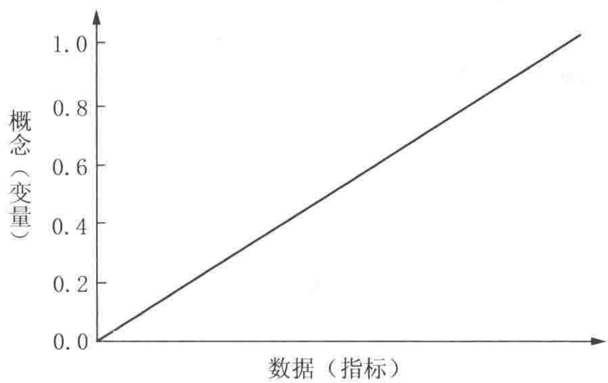


图 11.1 语义转换:线性

图 11.1 描绘了一种图解数据和概念之间关系的方式。其中 X 轴为“数据”,Y 轴为“概念”。这个图可以恰当地说明模糊集分析,因为研究者想要知道,X 轴的数据如何与 Y 轴表示的概念的语义意义联系起来的。相比之下,此类图并不用来汇总统计分析中的指标和概念之间的关系。图 11.1 假定这种关系是完全线性的。当定量研究者不依据初始量纲来转换数据时,这就是他们给出的模糊逻辑解释。然而,在模糊集分析中,几乎从来不做这种从数据到概念的线性变换。

反之,当将数据变换为隶属值的时候,语义学的模糊集方法考虑的就是概念的意义。例如,当将“高”这个概念应用于人的时候,研究者会

① 在一些社会科学以及在某些时期内,人们对标准化的变量有一种偏爱。尽管变量的标准化不改变实质的结果,但是它的确允许变量之间进行对比,因为它们都有相同的单位,即标准差。

考虑“高”的含义(Zadeh，1965)。在探究“个子高的人”的语义时，一个简单的方式是追问身高为六尺三寸(1.9 米)的人是不是高个子。一个可能的回答是，“是的，他当然是高个子”。转化成模糊集术语的话，此人在高个子集合中就有 1.0 的隶属值。对于一个五尺七寸(1.7 米)的人来说，可以问同样的问题。回答可能是他当然“不高”，这意味着他在高个子集合中的隶属值是 0.0。那么，6 尺(1.83 米)高的人呢？这样的人是“某种程度上的”高个子。这个人在高个子集合中可能有 0.50 的隶属值。这些“半空、半满”的点在语义学变换中是临界点。

0.50 这个隶属值大致类似于中位数或均值，因为它代表一个中间点。然而，实际上，0.50 隶属值的运作方式完全不同于统计均值。在模糊集逻辑中，0.50 隶属值是一个“交叉”点：诸多个案从较大程度上属于集合到较少程度属于集合要经过这个点(反之亦然)。它也是这样的区域，即其中数据的小变化可能意味着大的概念差异(参见下文)。0.50 隶属值与数据的分布无关。大体上，人们在模糊集分析中几乎从来不将数据的均值作为 0.50 水平。例如，虽然一个六尺高的人在高个子集合中可能有 0.50 的隶属值，但此人在平均身高之上。

11.3 次要变异原理

图 11.1 呈现了一个线性的语义转换。事实上，在许多统计分析中将指标与潜在的变量联系起来的时候，这种线性变换是默认的步骤。在本节中，我们考量在定性研究中将数据与概念连接起来的一些常见的非线性语义转换。

我们针对“高个子”构造的隶属函数的方式可作为一个例子。我们说一个六尺三寸(1.9 米)的人完全隶属于高个子集合。那么七尺(2.1 米)高的人呢？此人显然会被认为完全是该集合中的一个成员。事实上，任何超过六尺三寸的人都是一个完全的成员。当然，同样的分析也适用于低于五尺七寸(1.73 米)的人：他们在高个子集合中的隶属值为

0。这意味着，高于六尺三寸或低于五尺七寸的变异相对于高个子概念来说没有语义学的意义或重要性。

现在，我们可以陈述在定性传承中的一个基本的语义转换原理：

次要变异原理：在数据中存在着拥有相同语义学意义的一些区域。

至于个案在多大程度上隶属于语义学范畴，数据中的变异不会总能转译为这种隶属程度的差异。不同身高的两个人可能都完全属于高个子范畴。

图 11.2 利用经济发达国家这个概念和关于人均 GDP 的数据提供了一个例子(参见 Ragin，2000 给出的类似的例子)。人均 GDP 上的差异(比如瑞士和瑞典之间的差异)并没有体现在它们隶属于经济发达国家集合的程度差异上；它们的隶属度都是 1.00。类似的是，大多数落后国家之间的差异都不能改变它们在经济发达国家集合中的 0 隶属度。

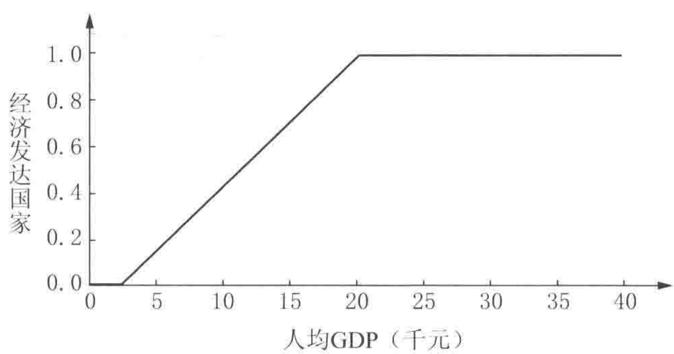


图 11.2　语义转换：次要变异原理

或许有人问图 11.2 中的那些隶属分值开始上升和下降的点。这些分界点不是“任意的”吗？一个合理的回答可能是，当研究者、政治家、著名的政府间组织(如联合国和世界银行)提到经济发达国家的时候，他们必定言有所指。如果他们确实言有所指，那么原则上人们可以

解码其意义，并用它作为基础来决定数据从哪里开始反映重要的语义变异。因此，有关这些分界点的选择无须是任意的。另外，只要直线的斜率是适度的（尽管斜率常常相对比较陡），隶属函数到底从哪个点开始下降可能就不那么紧要了。

关键点是相当直观性的：在顶部的 4 000—5 000 美元的差异根本不要紧，但是同样的差异出现在中部可能就非常重要。这导致了次要变异原理（Principle of Unimportant Variation）这个推论，也可称之为小异大效原理（the Principle of small differences but big impacts）。人均 GDP 的差异对于经济发达国家集合的成员来说无关紧要，这个差异有一个范围。与此范围类似，还有被这些差异放大的范围。放大的差异出现在小于 1 和大于 0 的隶属值之间。在这里，人均 GDP 上的小差异被转化成经济发达国家集合的隶属值中的大差异。

在图 11.2 中，虽然直线的斜率对于 Y 轴上所有大于 0 且小于 1 的隶属值来说是恒定的，但它不必用模糊集分析这种方式进行精确地转换。在定量传承中，如果默认的语义变换是线性的，那么在模糊集分析中最受欢迎的选择就是 S 曲线的某种变体（再次参见第 12 章）。依据确切的 S 曲线，可以在不同的地方发现放大的影响域，尽管它通常围绕在 0.50 隶属值周围。简而言之，次要变异原理通常运用于数据范围顶端和底端，隶属值是 0 或 1。相比之下，变动的效应在 0.50 隶属水平附近被放大，此处数据范围内的小改变常常对应着大的语义差异。

11.4 隶属函数和量表类型

假定读者在方法论课程上学习过典型的尺度类型分层。它们按顺序是：定类、定序、定距及定比的尺度。要想描述这些尺度之间的差异，通常要部分基于它们在比较个案时所包含的信息量。例如，定类尺度包含的信息较少，因为它们只可用于类别比较，定距和定比尺度包含的信息量最多，因为它们可用于进行规模测量。自史蒂文斯发表经典文

章(Stevens, 1946, 1968)以来,这些尺度类型已成为任何基础统计学训练的一部分。

那么,这些尺度类型是如何与模糊集分析中的隶属函数相匹配或不匹配的呢? 一个好的起点是零点。定距尺度之所以不同于定比尺度,是因为后者有一个真实的零点,而前者没有。例如,你的银行账户中的钱是一个定比尺度,因为达到零点就会使账户清零,欠债与余额为正的差异巨大。

在模糊集分析中,0 隶属值却不扮演同样的角色。在模糊集分析中,0 值表明完全不是一个集合中的成员。例如,一个国家的非零的人均 GDP 是 500 美元,它在"富裕国家"类别中的隶属值无疑是 0.0。从这个意义上看,模糊集分析中的 0 值接近于一个连续量表上的最小点思想,尽管二者也有差异。

针对模糊集量表中的 1.00 隶属值也有同样的论点。1 值表明完全隶属于一个集合。它扮演的角色在某种程度上类似于一个连续量表中的最大点。然而,在经典的量表架构中并不存在真最大值这个观念。例如,说一个人的富裕程度最大,这是什么意思呢? 一个人总得有一元钱吧。

经典的测量理论将两分变量看做占据着尺度层次的低端。如果有连续编码的数据,那么将它转换成两分数据几乎总被认为是糟糕的转换,因为这会丢失信息。模糊集分析提供了一个非常不同的二值化观点。我们可以通过考察通常用于测量民主的政体数据(Polity data)(Marshall, Gurr and Jaggers, 2010)来展示这个不同。

图 11.3 使用了学者运用于政体数据的一个常见的二值化架构,其中 7 是−10 到 10 的民主量表上的分界点。从统计学立场看,对政体数据进行二值化会丢失大量的信息:我们从一个 21 点的量表变成了 2 点的量表(参见图 11.3)。相比之下,从模糊集的观点看,二值化只是另一种语义变换函数。模糊集分析反对这类二值化,不是因为它丢失信息,相反,反对的是曲线的斜率,即在 7 这个点的垂直线太陡了(其导数是无穷大)。这里有一个分界点的确重要的情况:数据的微小改变可能导致语义意义的完全改变。

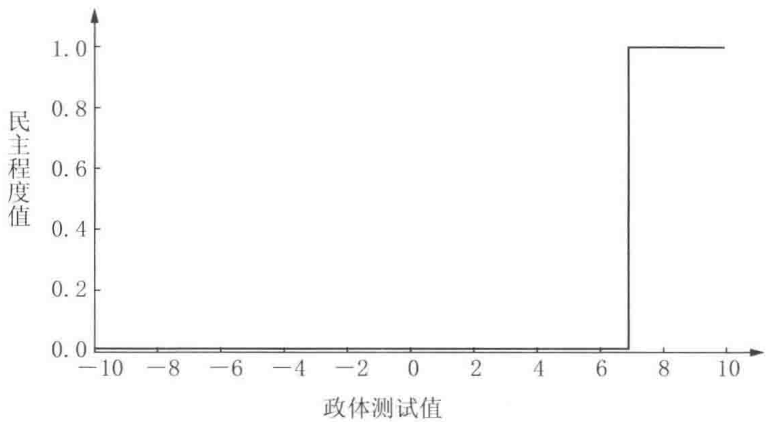


图 11.3　二分变量和成员资格函数

然而，如果调整斜率，使之不如此极端，那么对于定性学者来说，在评估一个国家隶属于民主国家集合的程度方面，二值化步骤作为一种方法看起来是合理的。现在，个案的模糊集隶属值会随着政体得分(Polity scores)降到分界点 7 以下而从 1 开始逐渐下降。可以通过比较图 11.2 和图 11.3 中的变换(直线表示的部分)看到这一点。虽然图 11.3 是二值化的，但是其基本形式实际上与图 11.2 差异不大。差异仅仅是连接最大(即 1.0)和最小(即 0.0)隶属值的直线的倾斜程度不同罢了。对于模糊集分析者来说，二值化的隶属函数虽然有问题，却可能比隐含在 21 点政体得分中使用的默认的线性假设更恰当。

11.5　结论

在本章中，我们的分析表明，模糊集方法可以很好地描述定性研究者通过什么方式来考量数据和概念隶属之间的关系。在这种方法中，一个关键的问题涉及通过恰当的语义变换将数据转换为隶属值。研究者关注将数据与概念的意义进行匹配，因此，要假定概念的意义支配了适当的变换。在此意义上，就模糊集分析来说，概念化的意义就统领

全局。

在定量方面，模糊集分析中使用的非线性变换的类型却不经常用于连接数据和变量。相反，由于数据被作为一个变量的指标来使用，因此，通常并且恰当的看法是将数据视为处在与变量的线性关系中。一个统计变量的好指标一般不要求变换。相反，指标值会与变量值紧紧匹配。就此而言，在统计分析中，我们应该寻找那些在测量中不要求变换的指标。

参考文献与推荐读物

对于本章描述的语义学研究来说，关于概念的定性文献可以说是一个好的切入点。参见科利尔和耶林(Collier and Gerring，2009)；科利尔和列维茨基(Collier and Levitsky，1997)；格尔茨(Goertz，2005，2008)和萨托里(Sartori，1970，1984)。有关统计研究中常见的变量-指标研究，参见布莱洛克(Blalock，1982)和博伦(Bollen，1989)。除了少数的例外，如阿德科克和科利尔(Adcock and Collier，2001)以及格尔茨(Goertz，2005)，这两类文献之间很少有沟通。模糊集合与语义学之间的联系可追溯到扎德(Zadeh，1965)，并且在柯思科(Kosko，1993)，麦克尼尔和弗赖贝格尔(McNeill and Freiberger，1994)的优秀导读性文献中有所强调。拉金(Ragin，2000)在利用人均 GDP 作为例子的研究中讨论了次要变异观念。

Adcock，R.，and D. Collier. 2001. "Measurement Validity：a Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95:529—546.

Blalock，H. 1982. *Conceptualization and Measurement in the Social Sciences*. Beverly Hills：Sage Publications.

Bollen，K. 1989. *Structural Equations With Latent Variables*. New York：John Wiley & Sons.

Boyd，L.，and G. Iverson. 1979. *Contextual Analysis：Concepts and Statistical Techniques*. Belmont：Wadsworth.

Collier，D.，and J. Gerring，eds. 2009. *Concepts and Methods in*

the Social Sciences: The Tradition of Giovanni Sartori. London: Routledge.

Collier, D., and S. Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49: 430—451.

Gates, S., H. Hegre, M. Jones, and H. Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800—2000." *American Journal of Political Science* 50:893—908.

Goertz, G. 2005. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.

Goertz, G. 2008. "A Checklist for Constructing, Evaluating, and Using Concepts or Quantitative Measures." In *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmier, H. Brady, and D.Collier. Oxford: Oxford University Press.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Kosko, B. 1993. *Fuzzy Thinking: The New Science of Fuzzy Logic*. New York: Hyperion.

Marshall, M. G., T.R. Gurr, and K. Jaggers. 2010. *Polity IV Project: Political Regime Characteristics and Transitions, 1800—2009*. Center for Systemic Peace, www.systemicpeace.org/polity/polity4.htm.

McNeill, D., and P. Freiberger. 1994. *Fuzzy Logic*. New York: Simon and Schuster.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Sartori, G. 1970. "Concept Misformation in Comparative

Politics." *American Political Science Review* 64:1033—1053.

Sartori, G. 1984. "Guidelines for Concept Analysis." In *Social Science Concepts: A Systematic Analysis*, edited by G. Sartori. Beverly Hills: Sage Publications.

Stevens, S. 1946. "On the Theory of Scales of Measurement." *Science* 103:677—680.

Stevens, S. 1968. "Measurement, Statistics and the Schemapiric View." *Science* 161:849—856.

Zadeh, L. 1965. "Fuzzy Sets." *Information and Control* 8: 338—353.

12 语义学、统计学和数据转换

数据转换(如幂次)可以使一个偏态分布变得比较对称。但是,我们为什么还要困扰呢?

(1) 高度偏态的分布难以考察,因为其大多数观测值都局限在小部分数据范围中。

(2) 显然,当把偏态分布变成比较对称的分布时,偏态方向中的离群值会变得接近于数据中的大部分值。

(3) 一些常见的统计方法用均值对分布进行汇总。然而,一个偏态分布的均值并不是它的中心的一个好的汇总值。

——约翰·福克斯(John Fox)

12.1 导言

在统计学传承中,有关变量的转换存在着一些公认的规范,它们使得对数据实施诸如标准化和求对数这样的做法变得很常见且无争议。在已知这种传承的研究目标的情况下,这些做法在方法论上相当易于理解。然而,定性角度则强调理解个案的意义,从这个角度看,同样的做法就会出现问题。定量传承中的变量转换回应了统计学的律令;定性学者则运用一套不同的规范和价值,它们强调语义学的重要性以及概念中体现的意义。例如,对人均 GDP 求对数或对民主的政体测度值进行标准化意味着什么,这种传承提供了备选的解释。

在本章中，我们介绍变量转换的基本原则（Fundamental Principle of Variable Transformation），将它作为描述定性角度下的变量转化的一种方式。该原则认为，所有的变量转换都必须使其意义得到保留或增加。因此，该原则要求，如果使用人均 GDP 的对数值，得到的数据应该比未转换的数据更好地代表学者所意指的概念，如经济发展或财富。在定性传承中，与这条原理不符的转化被认为是可疑的。

12.2 标准化比对意义保留

常见的转换是将一个变量标准化。在统计分析中，标准化通常不改变结果，因为对变量执行线性变换时，[①]大多数参数估计值仍然会保留其特性，如无偏性。回想一下，对变量执行标准化时用到的公式是 $(x_i - \bar{x})/s$，其中 $\bar{x}$ 是均值，s 是标准差。例如，初始的民主政体量表将数据编码为从 −10 到 10（Marshall，Gurr and Jaggers，2010）。标准化的过程就是将这些数字转变为一个新的量表，其值表示远离均值多少个标准差。

标准化的结果显然依赖于基础数据，而不依赖于与所测量的概念定义有关的数据。统计均值和标准差会随着基础数据的改变而变化。事实上，就标准化来说，如果其他个案的得分改变，那么任何一个个案的得分都会轻易地改变。一个个案可能没有经历任何真实世界的变化，但是其编码却可能改变了，因为其他个案的编码有变。从定性视角来看，这看起来很奇怪：一个个案的得分为什么应该依赖于如何对其他个案编码呢？

为了便于具体论述，我们考察关于民主程度的政体数据集。定性研究者可能提出一个规则，认为“完全民主”刻画的是那些超过一定阈值的

---

① 然而，如果进一步转换或分析变量，情况就可能有变。例如，如果被转换的变量用在交互项时即如此。

个案，如所有其值在 7 到 10 之间的个案。在该规则之下，任何一部分个案可能全是或全不是完全民主的国家，这取决于是否达到阈值。然而，如果数据是用标准化的值来展示的，那么人们必须使用不同的规则，例如，所谓完全民主描述的是那样的个案，即它们的值超出均值两个标准差。该规则确保仅有少部分个案被编码为完全民主。一个既定的个案可以移进或移出完全民主类别，这取决于其他个案怎样被编码。从定性视角来看，这是有问题的。一个个案是否完全民主，这取决于该概念的定义和该个案的特征，不取决于作为一个整体的总体内部的民主程度分布。

标准值的使用对于研究来说也可能有其他后果。例如，可以考虑这样的建议，即人们应该对变量实施标准化，以便基于其"极端"值来选择个案研究(Gerring and Seawright，2008)。至于人们为什么要考察极端个案，则有一些好的理由。鉴于此，标准化的值就是一个显而易见的选择，因为从标准差的角度看，我们认识到极端个案意味着什么：距离均值 2—3 个标准差的观测值就是极端值。因此，从统计传承内部看，这种思路完全可以理解。

然而在定性传承中，运用这种思路就遇到了问题。标准化思路是依据取值与样本均值的关系来界定极值的，均值被看成是"中间点"。但是，对于定性研究者而言，样本均值可能代表或不能代表一个概念的中间点。中间点对应着概念的中间值。例如，政体量表的范围是从－10 到 10，其概念的中间点可能是 0。[①]相比之下，政体数据的样本均值大约是 3。

在政体数据中，标准化的一个后果是大多数极端个案都总是威权政权。由于样本均值大约为＋3，因此，－10 的威权个案总是比＋10 的完全民主个案有更大的绝对标准值。在统计传承中，这一点未必有问题：大多数威权个案从它们与均值的偏差角度来看都是比较极端的。然而，从定性的角度看，标准化会适得其反。从这一角度看，极值"明显"是－10 和 10 的个案。这些个案就与量表的概念中间点的关系方面，最可能存在极值。

① 政权类型的"政体测度值"等于民主减去威权，其中民主和威权的范围都是从 0 到 10。在这种情况下，0 在"政体"计算程序中就是概念的中间点。

这就引出我们在定性传承中所说的变量转换的基本原则：

> 变量转换的基本原则。所有的数据转换都应该强化数据和概念的意义之间的关系。

在前一章中，我们使用了有关“高个子”概念和高度数据的例子。定性研究者可以转换高度数据，但是其动机是要实现与所关注的概念更匹配。相比之下，标准化（如对高度数据标准化）的动机通常不是与任一概念相匹配，因此，它就违反了变量转换的基本原则。

简而言之，定性的、语义学探究通常将标准化视为一步倒退。它将一些无关的考量（即潜在地变动着的数据分布）引入到一种语义关系之中。数据的标准化通常违反了变量转换的基本原则。它加入了现实世界的数据分布的某些方面，这些方面没有益处且有潜在的误导性。①

12.3 对数转换—模糊集转换

取一个变量的对数（即用其自然对数）非常常见，并且在统计研究中经常被推荐使用。举一个经典的例子，对人均 GDP 取对数很少有争议。虽然并不是所有的学者都运用这一转换，但我们仍然可以看到，在各类研究中有无数例子都使用了这种转换。

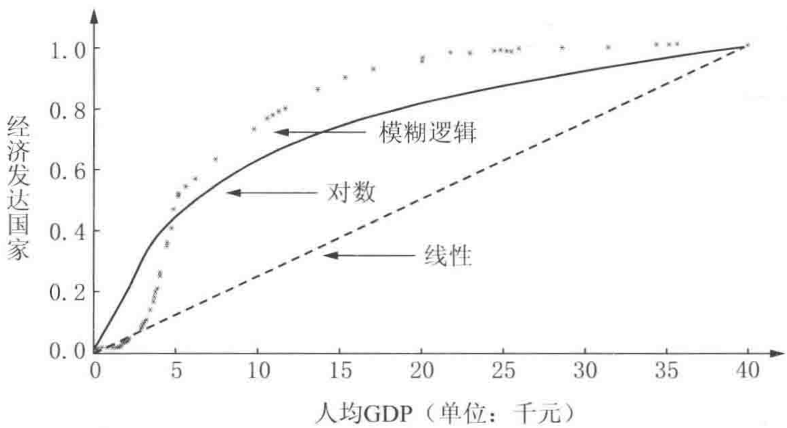
在定性方面，模糊集分析提供了一套工具，用它来设定人均 GDP 和财富或经济发达国家概念之间的关系。模糊逻辑关注的是，数据在多大程度上表明个案隶属于财富集合或经济发达国家集合。根据标准的惯例（参见第 11 章），得分为 0 意味着一个国家不属于财富集合或经济发达国家集合，得分为 1 意味着完全隶属，即个案是富裕的或经济发

---

① 至于为什么标准化经常不是一个好做法，弗里德曼（Freedman，2009：87—88）利用胡克定律（Hooke's law）（参见第 16 章，该章将胡克定律作为一个例子来考量）给出了精彩的展示。

达的国家。个案的隶属得分也可以在 0 和 1 之间。

在使用人均 GDP 去评估各个国家在多大程度上是富裕或经济发达的国家时，两种传承采取不同的方式，考察一下两种方式的差异是有益的。通常情况下，定量传承假定人均 GDP 水平和国家发达程度之间存在着线性关系，如图 12.1 中的虚线所示。这里通常有一个隐含的假定，即经济发展随着人均 GDP 线性增长。然而，如果使用对数值，我们就得到图 12.1 中的实线。对这条线有一个潜在的解释是（另一种解释见下文），在人均 GDP 和经济发展之间存在着收益递减的关系。换言之，图 12.1 中对数曲线的斜率随着人均 GDP 的增加而降低。



资料来源：Ragin，2008：92—93。

图 12.1　线性、对数和模糊逻辑的数据转换

在模糊集分析中，通常假设连续的定量数据和隶属得分之间的关系呈 S 形曲线。在图 12.1 中，人均 GDP 和经济发达国家之间即遵循这种模式，如图中的星线所示（见 Ragin，2008：第 5 章）。在这种模糊集的转换中，人均 GDP 低的个案都得到同样的 0 编码值；从语义学观点看，这些个案在所关注的概念（经济发达国家）中都没有隶属关系。这就是在第 11 章中所讨论的次要变异原理。这些个案处在人均 GDP 非常低的水平，接近图 12.1 的原点。

一旦个案开始隶属于一个概念，小的差异就会大大地影响它们在

多大程度上成为经济发达国家集合的成员。因此，尽管当人均 GDP 处在 2 500—5 000 美元时，模糊集曲线的斜率变得更陡，但是模糊集曲线总体来讲类似于对数曲线。从模糊集曲线的另一面来看，拥有高水平的一系列个案全被编码为 1，因为它们都是概念的完全成员。其结果是，模糊集曲线的右侧变得显著平缓下来。

就对数值来说，在人均 GDP 顶端的变动也变得相对不重要了（尽管斜率没有变成完全平缓）。低端的变动（例如在人均 GDP 低于 2 500 美元的个案之间的变动）却非常重要。这是与语义的、定性方法的根本差异。从语义学的角度来看隶属于"经济发达国家"的程度，可以认为低端个案之间的差异并不十分重要，因为这些国家显然不是经济发达国家。例如，乍得（Chad）和马里（Mali）的人均 GDP 之差不重要；它们隶属于经济发达国家类型的程度都是 0。但是对于使用人均 GDP 对数的学者来说，低端之间同样的这些小差异就会因对数转换而放大（如图 12.1 所示，对数线从原点开始就逐渐变陡）。从这个角度看，乍得和马里之间的人均 GDP 之差对于它们的经济发展水平来说就相当重要了。对于贫困的国家来说，人均 GDP 的小幅度增长意味着其经济发展的大幅度进步；而对于富裕国家来说，小变动的后果不大。

这一差异与两种传承的不同规范有关。在定性传承中，人们必须总是追问正在测量的特定概念的意义是什么，将它理解为各个个案所隶属的集合。研究者的责任是，将数据进行转化，使得人均 GDP 与研究者所意味的富裕或发达国家更好地匹配。相比之下，在定量传承中，研究者是因为数据偏态或要增加统计模型的拟合度才对数据求对数的。从定性视角来看，这些类型的考虑似乎并不恰当；而从统计学的立场出发，正如本章导语所表明，这些考虑很容易理解。

12.4 数据转换的基本原理

为了理解这些可替代的数据转换的基本原理（rationale），考察一

下不同水平的人均 GDP 个案的分布是有益的。如图 12.2 所示，大多数个案的人均 GDP 都很低。从模糊集的视角看，人们的结论是，大多数国家都不隶属于经济发达国家的范畴中。

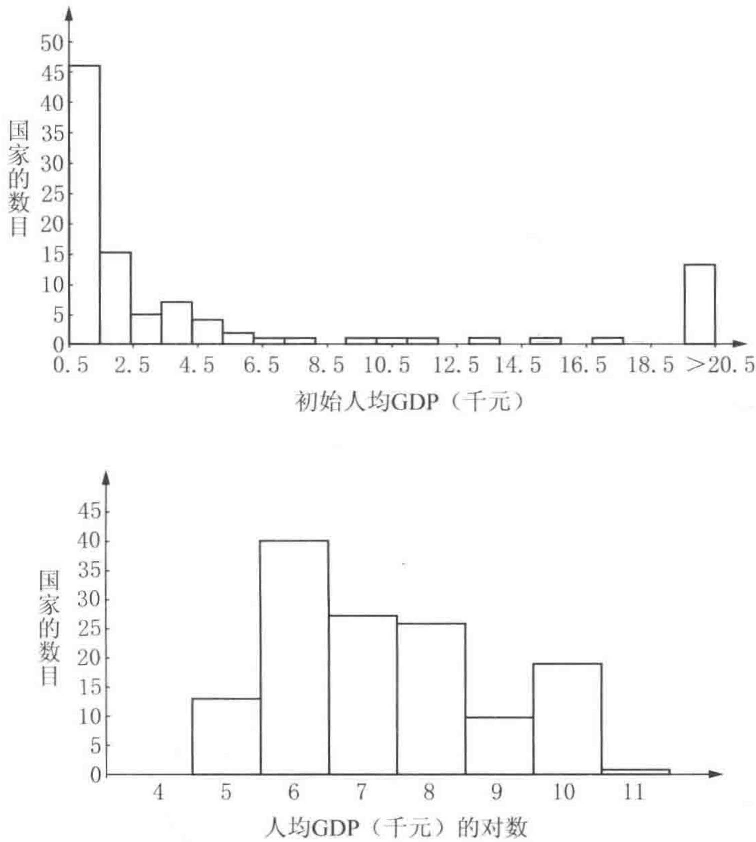


图 12.2　依据初始和对数化的人均 GDP 给出的国家分布

图 12.2 直方图也清楚地表明，有相当多的国家是非常富裕的（即人均 GDP 高于 2 万美元的个案）。这些个案是被编码为完全隶属于经济发达国家集合的国家。位于中间的个案数量（模糊集合中隶属值介于 0 和 1 之间）相对较少。多数个案要么完全属于，要么完全不属于经济发达国家。

在模糊集转换中，原始数据和切割点之间是什么关系？尤其是，针对完全属于、不属于、一定程度的部分隶属的个案，研究者如何决定在

哪里绘制截止点？在图 12.1 中，数据和编码决策源自拉金（Ragin，2008）。如何运用专业知识来绘制截止点，他的讨论为我们提供了一个好例子。在他的编码决策中，关键因素是确定完全隶属（人均 GDP 为 2 万美元）、无隶属（人均 GDP 为 2 500 美元）的阈值，以及类别成员和非类别成员（人均 GDP 为 5 000 美元）的个案数相等的交叉点，并在模糊集合中被编码为 0.5。他研究了一些学者和如世界银行这样的重要机构是如何确定一些国家是富有还是贫穷的，从而部分地获得了切割点。他在设计图 12.1 中 S 曲线的形状时使用了它们的语义学实践。因为学者和大多数机构的现实世界语言都支持转换，最终的个案隶属于集合的分值对于该共同体来说是有意义的。也就是说，国家的隶属值应该符合专家对财富和经济发展的概念定义和探讨。

因此，在定性方法上，人们在实施数据变换时，采取的方式应该旨在满足变量转换的基本原则。人们对原始数据进行转换，以便它们更好地匹配分析者或相关的共同体所意指的概念。由于它在一定程度上是一个解释的过程，分析者在进行数据转换时，理想的情况是清晰地阐述所使用的标准。这些标准可能反映着大众语言和普遍的文化习俗、专业知识及用法，和/或研究者关于所研究的现象的实质知识。用线性、取对数或标准化的转化，可能保留或提升一个既定概念的意义。但是这一点不能被先验地假定成立；它需要通过明确地评价个案的编码和概念的意义之间的匹配程度来展示。

现在转向统计传承中的对数转换，图 12.2 下方的图清晰地表明，当使用对数值时个案的分布产生了巨大变化。现在，数据接近于正态分布。通过使初始人均 GDP 很低的许多国家之间的差异变得明显，即可产生这种分布。因此，人均 GDP 少于 2 500 美元的国家构成了图 12.1 下面直方图左方的大多数变异。人均 GDP 低端的变异与语义学的、模糊集合分析不相关，这在对数化的数据集合中至关重要。

在如图 12.2 这样的例子中，当定性研究者看到取对数后的个案分布有变化时，他们可能怀疑这一方法。在此例子中，新数据意味着大多数国家处于中等的经济发展水平，定性研究者或许感到这不可信。他或她可能关心的是，变换歪曲了潜在的经验真实。

然而，从定量传承内部角度看，针对对数值的使用有相当多有道理的辩护。有时候，出于理论原因而使用对数值。例如，琼等(Jone et al., 2009)论证并从经验上指出，几乎所有的政府预算都遵循幂律分布。这种分布是 $y \sim ax^b$，它意味着明显的经验检验就是对等式的两边取对数。这种转换是合理的。

取对数的另一个常见的理由在于数据是偏态的。取对数在消除偏态上是有效的，如图 12.2 所示。这种转换背后的基本原理要求关注数据的真实世界分布问题。如果你的研究目标是用统计方法进行有效的推断，那么数据的偏态就是一个严重的问题，需要纠正。

最后，取对数的另一个道理是结果数据能够更好地拟合统计分析，所谓"更好地"意味着有更强的实质性影响或更高的显著性水平。这类道理可归入设定搜索(specification searches)这个标题之下(例如 Leamer, 1978)，该标题在过去几十年里引起了统计学家的广泛讨论。人们可以想象进行各类变量转换，然后选出一类能给出最强结果的变量转换。人们偶尔会发现，提升的拟合度是否辩护了一种特定类型的变量变换，定量学者是有争议的。①然而，从定性的、语义学观点看，模型拟合问题应该与如何最好地转换原始数据这个问题完全分开。仅仅为了提高统计拟合度而转换变量是冒风险的，即它可能违背变量转换的基本原则。

12.5 结论

定量和定性研究都对原始数据进行转换。然而，它们在实施这些

---

① 例如，库尔茨和施兰克(Kurtz and Schrank, 2007)认为，治理与经济发展无关。世界银行经济学家回应说："在下一组调查中，我们将展示与库尔茨和施兰克的初始设定有两项微小变化的后果。与他们不同，我们不加入人均 GDP 水平，而加入其对数水平值。这种做法在跨国的经验研究中是非常标准的做法，并且在统计上也比较恰当，因为因变量和对数的人均 GDP 之间的关系更接近于是线性关系，并且我们使用的是线性回归模型。"(Kaufman et al., 2007:59)。

转换时所遵守的规则大不相同。定量研究者采用诸如标准化和取对数值等做法帮助他们努力执行良好的统计检验。定性研究者几乎不对原始数据进行标准化和取对数。相反，他们“校准”(Ragin，2008)数据，使之符合分析者和(或)相关的专家共同体所界定的概念的意义。

需要再次指出的是，我们并不认为一种方法一定优于另一种。实际上，我们认为，两种方法在各自的总体传承的情景下都是可以理解的。因此，我们已经表明，在统计传承中，对原始数据进行标准化或执行对数转换都很有道理。同样，在定性传承中，在针对一个概念将隶属值指派给个案的时候，忽略一些变异也可以理解。在一种传承中可能是不合适的做法也许在另一种传承中相当合适。

参考文献与推荐读物

一些导读性的统计学教材讨论了如何计算标准值，如何进行对数转换以便消除偏态性。关于如何利用模糊集分析对初始的定量数据进行转换，尤其可以参见拉金(Ragin，2008)。有关定量和定性的转换方面的例子，参见下面有注释的文献。

Brown，J.，J. Earle，and S. Gehlbach. 2009. “Helping Hand or Grabbing Hand? State Bureaucracy and Privatization Effectiveness.” *American Political Science Review* 103：264—283. 在该例子中，大多数变量都取了对数。

Busse，M.，and C. Hefeker. 2007. “Political Risk，Institutions and Foreign Direct Investment.” *European Journal of Political Economy* 23：397—415. 他们想将“对外直接投资”变量取对数，但是那将导致负数，因此，他们利用 $y = \log(x + (x^2 + 1)^{0.5})$ 对数据进行了转换。

Carment，D.，S. Yiagadeesen，and S. Prest. 2009. “State Fragility and Implications for Aid Allocation：an Empirical Analysis.” *Conflict Management and Peace Science* 25：349—373. 注意他们如何用婴儿死亡率(infant mortality)的对数值和标准化值来表示生活标准的指标。

Fearon, J. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8:195—222.他认为,族群多样性的一个测度要好于另一个,因为个案的分布不是偏态的。

Fischer, C., M. Hout, M. Jankowski, S. Lucas, A. Swindler, and K. Voss. 1996. *Inequality by Design: Cracking the Bell Curve Myth*. Princeton: Princeton University Press.作者们批判了对钟形曲线(Bell Curve)数据实施正态化处理。

Freedman, D. 2009. *Statistical Models: Theory and Practice*, revised edition. Cambridge: Cambridge University Press.

Gerring, J., and J. Seawright. 2008. "Case-selection Techniques in Case Study Research: a Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61:294—308.

Goertz, G., and J. Mahoney. 2005. "Two-level Theories and Fuzzy-set Analysis." *Sociological Methods and Research* 33: 497—538.作者们将来自斯考切波(Skocpol)的《国家和社会革命》(*States and Social Revolutions*)的定性数据重构为模糊集合。

Herrnstein, R., and C. Murray. 1994. *Bell Curve*. New York: Free Press.他们使用的智商检测值不是正态分布的,因此,他们将这些值转换成正态分布;批评意见,参见 Fischer et al., 1996。

Jaggers, K., and T. Gurr. 1995. "Tracking Democracy's Third Wave with the Polity III Data." *Journal of Peace Research* 32: 469—482.

Jones, B., F. Baumgartner, C. Breunig, et al. 2009. "A General Empirical Law of Public Budgets: a Comparative Analysis." *American Journal of Political Science* 53:855—873.

Kaufmann, D., A. Kraay, and M. Mastruzzi. 2007. "Growth and Governance: a Reply." *Journal of Politics* 69:555—562.

Kurtz, M., and A. Schrank. 2007. "Growth and Governance: Models, Measures, and Mechanisms." *Journal of Politics* 69: 538—554.

Leamer, L. 1978. *Specification Searches: Ad Hoc Inferences With Nonexperimental Data*. New York: John Wiley & Sons.

Londregan, J., and K.Poole. 1996. "Does High Income Promote Democracy?" *World Politics* 49:1—31.为了能够利用他们偏爱的统计方法,作者们利用下面的公式对"政体量表"(Polity scale)实施转换: $T(S)=\log(S+10.5)-\log(10.5-S)$,其中 S 是政体民主得分。

Mahoney, J. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. New York: Cambridge.作者利用定性历史数据来建构个案的模糊集成员值(fuzzy-set membership values)。

Marshall, M. G., T.R. Gurr, and K. Jaggers. 2010. *Polity IV Project: Political Regime Characteristics and Transitions, 1800—2009*. Center for Systemic Peace, www.systemicpeace.org/polity/polity4.htm.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.参见第5章对模糊逻辑变量转换和校准的详尽讨论。

Zadeh, L. 1965. "Fuzzy Sets." *Information and Control* 8: 338—353.

13 概念的对立面与类型学

二值逻辑并不是充当人文科学基础的正确逻辑。为此目的，我们需要的是模糊逻辑。本质上讲，模糊逻辑就是拥有不明显边界的分类逻辑。

——洛蒂·A. 扎德(Loti A. Zadeh)

13.1 导言

概念的一个奇特之处是可以用专业术语指明一个既定概念的对立面。概念常常以反义词对(opposing pairs)的方式出现，如民主-威权、战争-和平和富裕-贫穷。以一对形式出现的一个概念是否等价于另一个概念的反面，这是一个具有重要理论和方法论含义的核心问题。例如，是否富有＝不贫穷，战争＝不和平？我们可以更一般地提出如下问题：

> 概念的否等于其对立面吗？

用我们通篇使用的民主例子来说，不民主等于威权吗？

我们认为，在回答这个问题时，定性传统中的学者通常会说，“不”。定性学者采用一种非对称的方法，认为不民主是一个不同于威权的概念。相比之下，定量传统中的研究者对该问题的回答通常为“是”(至少

在方法论实践中)。当在统计模型中测量和使用概念时,定量传统接受一种对称的概念观及其反面。例如,可以将一个政体的民主程度理解为恰恰是其威权程度的否。

一个相关的问题涉及类型学中概念之间的边界。传统观点认为,类型学应该建立在互斥范畴的基础之上。每一个观测只能属于一个范畴。①一个国家不能同时被编码为"民主"与"威权"。然而在实践中,这种类型学分析通常与我们在世界中看到的不同。在政府的人口普查中,族群范畴就是一个明显的例子。我们也许会说,奥巴马在如下三个族群中拥有非零的成员资格:(1)非洲裔美国人;(2) 白人;(3)非洲人。然而,我们如果允许奥巴马总统在这三个族群中都有一定的成员资格,就违反了互斥类型学规则。②

在定性研究中,学者(在实际的研究实践中)总是反对互斥的类型学假定。在这种语义导向的传统中,说奥巴马总统至少在两三个不同的族群类型中享有部分成员资格,这是合乎情理的。在这个主题上,我们表明,有了这种非互斥的假定,模糊集分析自然是分析类型学的数学方法。

13.2 针对概念对立面的对称及非对称的研究

为了阐明针对概念对立面的不同研究路径,让我们考察民主和威权的概念。现在,为简便起见,想象用政体数据二分地测量民主概念(Marshall, Jaggers and Gurr, 2010)。政体量表的范围是−10 到 10,一个可能的二分动作是将得分在 7—10 之间的个案看成是民主的个案。那么,关键问题是怎么界定得分在−10 到 6 之间的那些个案?应该称这些个案为"威权",还是"非民主"的例子呢?

① 除此之外,一个类型通常应该在总体上是穷尽的,即全部观测值都应该被归类。

② 2010 年 8 月,格尔茨收到一封来自大学的电子邮件,信中说新的联邦条例允许个体在联邦表中选择超过一个族群。

在定量传承中，如果采用二分范畴（它通常不是首选），非民主和威权就是相同的。所有得分在－10 到 6 的例子都既是非民主又是威权的。这些个案完全有别于得分为 7 或更高的民主例子。为了阐明这种方式下二分范畴的排他性，图 13.1 展示了政体得分和一个个案隶属于民主集合的程度之间的关系。X 轴是民主程度测度，Y 轴是民主集合中的成员资格测度。正如该表所示，7 这个切割点可以用全有或全无式（all or nothing terms）来理解：在该点及以上，一个个案完全在民主集合中，在该点以下，个案完全在民主集合之外。

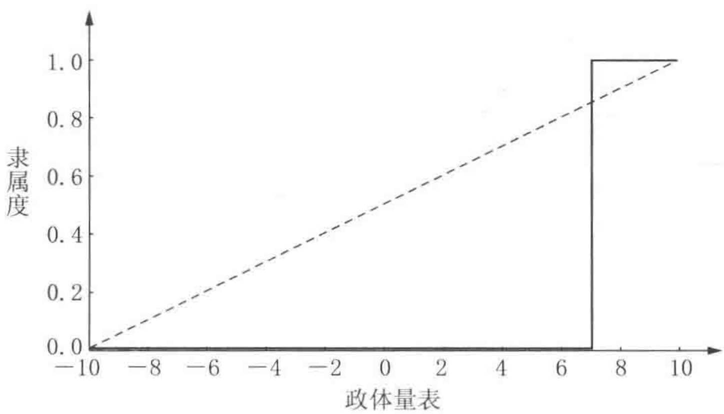


图 13.1　定量传承中对称的成员资格函数图式

当在定量传统中应用连续变量时，概念对立面的对称假定也适用（总体上讲，在定量传统中首选连续变量）。对于有关威权和民主的假设来说，通常选用同样的连续变量，这也是恰当的。一个个案在多大程度上是民主的，这不过是它在多大程度上为威权的否。在图 13.1 中，虚线展示了连续测量的政体得分和连续测量的民主集合的成员资格值之间的标准的线性理解的关系（参照第 11 章）。在定量传统中，同样的线性关系也适用于威权主义，除了它现在是逆向关系之外（即在威权主义中，政体得分与成员资格呈负向线性关系）。这种对称模式适用于很多概念。例如，经济发达与不发达都可以用同样的人均 GDP 指标来测量。如果有一个好的经济发展水平的测度，就有一个好的经济不发达水平的测度。

相比之下，这种对称性在定性传统中通常不适用。实际上，图13.1给出的政体测度和民主集合成员值之间的任何一种关系都不是定性传统中的一项好的选择。相反，当人们主要关注语义学和意义时，图13.2阐明了一种更典型的方法。有人也许论证到，只有当一个个案的政体得分为10的时候，它才在民主集合中拥有百分之百的成员资格(Goertz, 2008)。在文献中，二分标准将得分低于7的个案标记为完全不民主。然而，一旦有人从模糊集的角度来看民主，那么从较低水平——如图13.2中的4——开始看就是合理的。因此，得分在4和10之间的个案就部分是民主集合的成员，部分是非民主集合的成员。①

在模糊集分析中，X 的编码是从1到0，X 的否是1减去 X 的成员资格得分。因此，在图13.2中，虚线代表的非民主值等于1减去民主值(即 $\neg X=1-X$)。模糊集分析的否就是其字面意义：初始成员资格值的否。

然而，非民主和威权不是相同的概念。具体来说，它们不具有相同的成员资格函数。例如，人们采取一定方式将政体数据和威权联系起来，这种方式可能非常不同于将政体数据和民主联系的方式。例如，政治得分为−10到−4的个案或许被认为完全是威权类别中的成员；得分大于−4小于2的例子可能被视为在该类别中有部分成员资格；得分为2或更高的个案则在该类型中不具备成员资格。

关键在于，在这个传统中，民主和威权的概念并不对称。它们是不同的概念，因此有不同的成员资格函数。当然，非民主和威权有关，它应该如此。然而，它们不是相同的事物。从语义学的观点看，这一点对于很多对立的概念来说都成立。例如，大多数人都同意“不战争”与“和平”是不同的概念。以色列和埃及处在不战状态；但是并不清楚它们是否和平共处。

① 政体概念针对民主和威权有独立的量表。所以，尽管大部分学者通过从民主变量中减去威权变量来计算政体量表，但人们也可以把民主和威权分开，并给予不同的编码。

简而言之，由于一个概念的否在意义上不同于此概念的对立面，因此，语义学的路数会假定对一个概念及其反面的测量也不对称。每一个概念都要求从自身出发，依据它们自己的界定和意义来测量。我们称之为定性研究中的概念对立面原则（the Principle of Conceptual Opposites）。

> 概念对立面原则：一个概念及其对立面的意义和测量是不对称的。

13.3 重叠对比排他的类型学

定性研究者常常反对这样的观点，即类型学中的范畴必须互斥。例如，在图 13.2 中，对民主和不民主的模糊集编码明显违反了排他类别原则。一些个案（即政治得分大于 4 但小于 10 的个案）同时是民主和非民主类别的成员。同样的不排他性也适用于有三个及多个类别的类型学，如民主-半民主-威权，或者下层阶级-中产阶级-上层阶级。一些个案会同时属于多个类别。

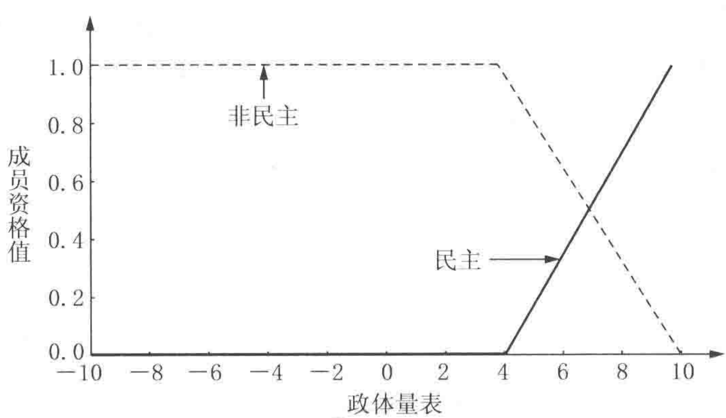


图 13.2 定性传统中民主及其否的概念化

继续用民主例子，很多学者最关注介于民主和威权中间的类型，有时候被定义为半民主(anocracy)。[①]当对这些个案实施编码的时候，在统计文献中使用的一种程序是将政体量表的中间领域(例如－5到5)当做半民主状态，将两端当做非半民主(即威权和民主)。这种办法已经应用于费伦和莱廷(Fearon and Laitin, 2003)关于内战的有影响力的研究中，现在这种方法已经普及。运用这种三分法的学者们通常将二分变量分成三类。图13.3展示了针对虚拟变量的标准三分编码。每个国家-年(country-year)都落入三类中的一类，且只落入一类。

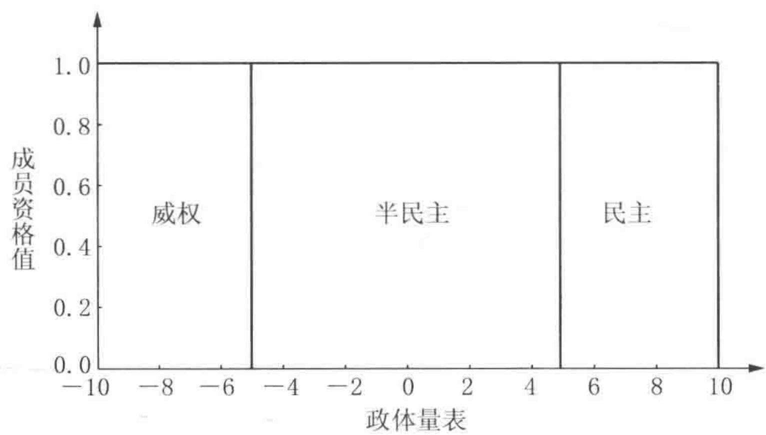


图13.3　民主-无政府-威权:互斥类

利用政体数据对这些概念进行模糊集编码会与图13.3中的编码有很大差异。模糊集分析通常不使用明显的拐点;模糊集方法也不遵循互斥分类规则。类别之间的过渡是逐渐的和部分的，而非突然中断。如图13.4所示，威权、半民主和民主之间可以有或多或少的重叠。某些国家可以同时处在威权和半民主状态，某些国家同时处在半民主和民主状态。

① 关于混合政体的大量文献(如Schedler, 2002; Lecitsky and Way, 2010)主要考察这种中间类型。中间类型对于关于内战(如Vreeland, 2008)和国际战争(Goemans, 2000)方面的研究来说也很重要。

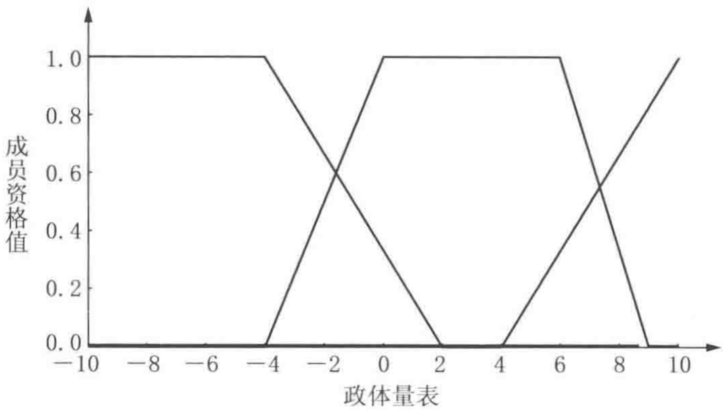


图 13.4　重叠类：民主、半民主、威权

为什么在统计分析中不使用这些模糊集编码呢？这里没有内在的原因。在统计学架构内，它们只是三个变量，其范围是从 0 到 1。然而，更常见的方法是图 13.3 中的互斥设置，这也是很多定性研究者所使用的方法。这本身没有什么不对，倒是反映出一种不同的决策，它决定如何划定类别之间的边界线。模糊集分析优先选择渐变过渡，互斥架构则暗含着突变。

总之，尽管定量研究中的类型学通常被理解为是由互斥的类别组成的，但语义学的概念研究则诉诸一个与之对立的观点。我们将这种定性的替代选择称为概念重叠原则(Principle of Conceptual Overlap)。

概念重叠原则：类型学中相邻的范畴可以重叠而非互斥。

无论在社会科学中使用模糊集合分析，还是在现实世界中用它来设计常用的器械时，概念重叠原则都居于核心。

13.4　语义学和名义类型学

考察在刻画政治制度中使用的概念的历史是很有趣的。表 13.1

重新展示了格尔(Gurr, 1974)最先提出的政治体制三分法,给出了三种名义类别。这些初始的类型是独裁政体、民主政治和无政府状态,它们各自都有由界定性特质构成的组合。可以推测,格尔认为这些范畴是互斥的。然而,他使用了至少五个定义维度(即表 13.1 中的威权变量),每个又至少有两个可能取值,即便如此,这三类显然仍然无法穷尽政治体制的全部可能领域。某些政治体可能拥有格尔的类型中的至少两个特性。

表 13.1 民主-半民主-威权:格尔(1974)的初始构想

权威变量	独裁政体	民主政治	无政府状态
行政部门聘用	归属=1 指派=1 双重=1	竞争性选举=2	恺撒式=1
决策限制	无限权威=2 轻微限制=1	立法平等=2 实质限制=1	无
参与	压制=2 约束=2	制度化的=2	非制度化的=2 派系/约束的=1
指导	威权主义=2 部分附加=1	无	最小化=2
集中化	无	去中心化=1	去中心化=2
最大分数	7	7	7

资料来源:Gurr, 1974。

如果目标是掌握有用的名义范畴,那么应该怎样处理不适用于已有范畴的那些个案呢?一种可能性就是简单地把它们当做"缺失数据",或者把它们排除在经验分析之外。然而这种办法看起来是有问题的,除非有明确的和可辩护的理由去放弃这些个案。一个替代的、语义学的解决办法是,询问这些个案是否可以被容纳到现存的类型学范畴内,或者它们是否要求有一个新的类别。这种语义学方法允许分析者在分析时将所有相关的个案囊括其中,要么将其放置在已有的范畴内,要么建立新的范畴。

为了进一步说明,我们再一次回到政体数据。这个数据因其对某些特殊个案的编码方式而声名远扬(或声名狼藉)。这些个案后来得到

了广泛的关注(如 Vreeland, 2008),因为在分析与内战和人权亵渎有关的结果时,它们影响了相关发现。下面有政体数据中三种特殊个案编码的官方声明(某些细节略):

> 中断阶段(—66):成分变量值"—66"代表着"中断阶段"。从操作角度讲,如果一个国家在战争期间被外部力量侵占,进而终结了旧的政治体制,那么政体数据库都将该个案编码为中断,直到一个独立政府得到重建。
>
> 空白阶段(—77):政体成分变量的一个"—77"编码值表征着"空白期",在此期间,中央政权完全倒塌。这最有可能发生在内战时期(例如,1978 年到 1986 年间的黎巴嫩)。
>
> 过渡阶段(—88):分值为"—88"表示过渡期。某些新政体之前有一个"过渡期",在此期间,新的制度正在筹划,合法地组建并付诸实施。在包含立宪会议和公民投票的程序中,民主的和准民主的政体尤其可能产生。在这个过渡期间,所有的指标都被记为"—88"。(Marshall and Jaggers, 2002:16)

在定量传承中,这些编码经常被看做一个缺失数据问题。一个标准的做法是将带有这些编码的个案排除在统计分析之外。在这种做法中,一些学者自然会试图估计"缺失值"。现在,政体数据库本身也有一个程序赋予"—77"和"—88"的个案以一定的政体得分。①

相比之下,对该问题的一个定性的、语义学研究始于考察在—66、—77 和—88 等编码中体现的概念。—66 的个案涉及"外国侵占",—77 的个案与"无政府状态"有关,—88 的个案则是那些"过渡期"类型。有人也许会问,这些类型个案的特点是怎样与民主、半民主和威权等概念匹配(或不匹配)的呢?一种可能性在于,它们仅仅是不同的概念,因而可能不适用于民主、半民主和威权清单。

① 普拉姆珀和诺伊迈尔(Plümper and Neumayer, 2010)批评这种政体数据做法,但是他们也假定该问题只是一种缺失值问题,这些个案应该被编码为—10 到 10 的政体得分。

无政府状态的－77个案是一个很好的开端。在初始的格尔编码中，无政府状态是一类常规类型。随着时间的推移，格尔的无政府状态对于某些个案来说变成了半民主，对于其他个案来说变成了“－77”。问题是，政体量表是政权类型的一种测度，而在根本没有正常运转的政府的情况下（例如，失灵的国家或者大规模内战的情形），我们应该如何探究个案，这是不清楚的。现在，政体数据在其量表中将这样的个案编码为0，但这似乎与中央政府缺失或垮台的半民主状态混淆了。由于“政府类型”不同于“政府存在”，从概念的角度来看，自然要考虑有一个新的变量，即“无政府状态”。无政府状态不是缺失数据的问题，它从概念上讲是单独一个问题。①对于很多经验研究项目来说，也许可以将相关的观测总体界定为所有拥有政府的国家。在这种总体的定义下，无政府状态的个案会被排除，但不是因为它们是缺失数据，而是因为它们不符合定义相关个案空间的标准。

－66是外强侵占的个案，它们也值得从概念的角度来思考。与无政府状态不同，在外强侵占下的国家是有政府的。大多数外国侵占者都是“军政权”。在第二次世界大战期间，德国侵占法国并没有造成无政府状态，而是产生了由德国人执掌的军事政府。因此，就概念角度而言，一种选择是把这些个案编码成像其他军事政权一样，即威权政体。它们中的有些个案也许是部分军事侵占的个案，就像美国在伊拉克和阿富汗一样。另一种选择或许会争论说，所有的国内政权都要求有一个自治和完全独立自主的国家。因此，军事入侵的个案就会落入一个独立的范畴，诸如“非主权政体”（non-sovereign polity），它可以包含所有的殖民个案。不管怎么样，这些外国入侵的个案都起源于概念问题，与缺失数据关系不大。

最后，相似的观点也适合于“－88”，即过渡期个案。这些个案的问题实际上并不在于缺失数据；人们关于缺失数据常常有丰富的信息。

① 很多人把内战当成是无政府状态的一种情境来分析（如 Malter，1997）。同样类型的问题也出现在国际冲突文献之中。例如，克莱因等（Klein et al.，2006）将两个国家之间的关系的实质概念化为：从友好到敌对。问题在于，当两个国家（如玻利维亚和缅甸）之间没有关系时，该怎么做？

问题在于，它们的政治制度仍然在规划进程之中，至于初期的机构应该如何与像民主这样的要求有清晰、明确和稳定的政治规则的类别匹配起来，这一点仍然不清楚。一个可能的途径是，考虑到这些过渡个案缺乏一个运转的政府结构，因而主张它们也属于“无政府状态”一类。一些人或许不同意，但这种不同意的本质涉及的是概念问题，而不是关于这些个案的经验特征的争论。

13.5 结论

定性研究者经常利用类别的反义对（如民主对威权）和类型学（如民主的、半民主的和威权的政权）。然而，他们的语义学方法却引导他们使用一些类别性的工具，使用的方式相当背离了定量研究的规范。他们未必将一个正概念及其反面看成是对称互逆的。一种关于发达的测度未必是不发达测度的否。因此，当定性研究者分析发达国家和不发达国家的时候，他们并不把它们看做互为映像的。对于定量研究者而言，这一点可能看起来奇怪，因为定量学者自然的默认选择是假定一个概念及其反面是完全对称的。在定量传承中，一个个案提高了发达水平，即可以理解为同时并且等量地降低了其不发达水平。

在实际的研究实践中，定性研究者有时也会拒绝一种相关的论点，即类型学分类必须是互斥的。他们经常允许邻近的类别之间可部分地相互重叠，因为这就是我们在日常语言中使用概念的方式。再次重申，从定量的观点看这也许相当奇怪，因为在这种传承中，类型学被看做由那些天然就互斥的名义范畴组成的。然而，从语义学的观点看，世界并不是恰好被分成完全分离的类别。在现实世界的许多个案所展示的类型学中往往有多个类别，很多个案同时隶属于多个类别，如果假定在名义量表中有明确的边界，就不足以捕捉到这种情况。

参考文献与推荐读物

定量学者对定性的类别和类型学的怀疑清晰地体现在金、基欧汉和维巴之中："诸如类型学、架构和所有的分类方式等构念都可用做收集数据的暂时的工具……。然而，一般来讲，我们估计研究者不要用这种方式来组织它们的数据。"(King, Keohane and Verba, 1994:48)其他定量的怀疑则建立在如下观念基础之上，即类型学会误导内在的变异，相关研究包括布莱洛克(Blalock, 1982:109)，杰克曼(Jackman, 1985:169)和夏夫利(Shively, 2005:32)。对这种怀疑的批判，参见科利尔、拉波特和西赖特(Collier, LaPorte and Seawright, 2012)。将名义量表看成是表达了互斥的类，这种理解可追溯到史蒂文斯的最初表述(Stevens, 1946)。在定性研究中，有关用互斥的类型学来描述和解释性分析的文献，参见科利尔、拉波特和西赖特(Collier, LaPorte and Seawright, 2008)，埃尔曼(Elman, 2005)，乔治和贝内特(George and Bennett, 2005)。在模糊集分析中，有关利用非互斥的类型学得到了良好的讨论，并主要关注工程上的应用，包括柯思科(Kosko, 1993)以及麦克尼尔和弗赖贝格尔(McNeill and Freiberger, 1994)的优秀概论。然而，这个观点在有关类型学的定性社会科学文献中并没有得到良好的发展。但是，在研究实践中，我们常常能够观测到非互斥的类型学的使用(例如，在关于混合政权的文献中)。

Blalock, H. 1982. *Conceptualization and Measurement in the Social Sciences*. Beverly Hills: Sage.

Collier, D., J. LaPorte, and J. Seawright. 2008. "Typologies: Forming Concepts and Creating Categorical Variables." In *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmier, H. Brady, and D. Collier. Oxford: Oxford University Press.

Collier, D., J. LaPorte, and J. Seawright. 2012. "Putting Typologies to Work: Concept Formation, Measurement, and Analytic Rigor." *Political Research Quarterly* 65:217—232.

Elman, C. 2005. "Explanatory Typologies in Qualitative Studies of International Politics." *International Organization* 59:293—326.

Fearon, J., and D.Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review* 97:75—90.

George, A.L., and A.Bennett. 2005. *Case Studies and Theory Development*. Cambridge: MIT Press.

Goemans, H. 2000. "Fighting for Survival: the Fate of Leaders and the Duration of War." *Journal of Conflict Resolution* 44:555—579.

Goertz, G. 2008. "A Checklist for Constructing, Evaluating, and Using Concepts or Quantitative Measures." In *The Oxford Handbook of Political Methodology*, edited by J.Box-Steffensmier, H.Brady, and D.Collier. Oxford: Oxford University Press.

Gurr, T. 1974. "Persistence and Change in Political Systems, 1800—1971."*American Political Science Review* 68:1482—1504.

Jackman, R.W. 1985. "Cross-national Statistical Research and the Study of Comparative Politics." *American Journal of Political Science* 29:161—182.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Klein, J., G. Goertz, and P. Diehl. 2006. "The New Rivalry Data Set: Procedures and Patterns." *Journal of Peace Research* 43:331—348.

Kosko, B. 1993. *Fuzzy Thinking: The New Science of Fuzzy Logic*. New York: Hyperion.

Levitsky, S., and L.A.Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. Cambridge: Cambridge University Press.

Marshall, M., and K.Jaggers. 2002. "Polity IV Project Dataset Users' Manual." Unpublished manuscript, University of Maryland.

Marshall, M.G., K. Jaggars, and T.R. Gurr. 2010. *Polity IV Project: Political Regime Characteristics and Transitions, 1800—2010*. Retrieved from http://www.systemicpeace.org/polity/polity4.htm.

McNeill, D., and P. Freiberger. 1994. *Fuzzy Logic*. New York: Simon and Schuster.

Plümper, T., and E. Neumayer. 2010. "The Level of Democracy during Interregnum Periods: recoding the Polity2 Score." *Political Analysis* 18:206—226.

Schedler, A. 2002. "Elections without Democracy: the Menu of Manipulation." *Journal of Democracy* 13:36—50.

Shively, W. P. 2005. *The Craft of Political Research*. Upper Saddle River, NJ: Prentice Hall.

Stevens, S. S. 1946. "On the Theory of Scales of Measurement." *Science* 103:677—680.

Vreeland, J. 2008. "The Effect of Political Regime on Civil War: Unpacking Anocracy." *Journal of Conflict Resolution* 52:401—425.

Walter, B. 1997. "The Critical Barrier to Civil War Settlement." *International Organization* 51:335—364.

第四部分　研究设计与普遍化

14 个案选择与假设检验

当基于因变量的一个特殊值来选择观测值时，如果不考虑到因变量取其他值的事例，我们无论如何也不会把握因变量的原因。

——加里·金(Gary King)

罗伯特·基欧汉(Robert·O. Keohane)

悉尼·维巴(Sidney Verba)

14.1 导言

有各种理由可以解释为什么学者会选择某些个案进行透彻的分析(Eckstein, 1975)。本章将探究的是，当分析的目的是评价因果假设的时候，在个案研究的选择方面有哪些定性和定量的实践。我们探索两种传承中使用的不同因果模型如何塑造了个案的类型，这些个案为假设检验提供了最强大的手段。对于检验关于平均处理效应的假设来说可能有一个好的个案，但是对于检验关于必要条件或者"充分非必要条件的必要非充分部分"(INUS conditions)的假设来说，它可能不是好个案。遗憾的是，关于个案选择的文献通常不关注所探究的因果模型的形式，或者仅假定调查者正在检验一个可加的线性统计模型。结果是，只有当我们的目标是估计平均处理效应而不是检验一个集合论假设时，文献中提供的大多数关于个案选择的建议才有意义。

我们将关于个案选择和假设检验的讨论分成两部分。第一部分

探索的问题我们认为在文献中最令人困惑：研究者应该基于个案在因变量上的值来选择个案吗？金、基欧汉和维巴（King，Keohan and Verba，1994）提议，定性研究者应该避免基于因变量来选择个案，这是一个重大的建议。然而，定性方法论者却认为，这个建议作为一个经典例子，恰恰说明了诸多洞见从统计研究不恰当地延伸到了定性研究（例如，Collier and Mahoney，1996）。我们会指出，当因果模型是（在一个大样本情境下）可加-线性模型而不是集合论模型时，这些建议怎样才有意义。在决定根据因变量来选择个案是不是一个好主意之前，研究者必须首先追问所研究的因果模型是什么类型的。

第二部分关注的是哪些类型的个案为个案研究中的因果推断提供了最大的影响力。这个问题在多元方法研究者中间产生了一类新的文献（如 Lieberman，2005；Seawright and Gerring，2008）。学者曾用个案研究的分析来补充他们的回归结果，现在他们通常坚持这类文献中的观点。然而，我们相信，对于想要检验集合论模型的定性研究者来说，此类文献中的有些建议是有误导性的。我们用一个简单的 2×2 表来论证此观点，并追问当进行个案研究时，每一种传承会优先考虑哪些格值。定性传承会优先选择(1，1)单元，在这个单元中，原因和结果都有呈现。这种传承也不愿意选择来自(0，0)单元的个案，因为这种单元对于评价集合论假设没有多大力度。相比之下，定量传承通常认为所有的单元都是重要的（Seawright，2002）。

14.2 基于因变量进行选择

定量学者为定性研究者提供的最生动、最著名的建议之一是，根据个案在因变量上的极值来选择个案是危险的。阿肯和斯尼达尔（Achen and Snidal，1989：160—161）将这种极端个案的选择——诸如乔治和斯莫克（George and Smoke，1974）决定主要关注战争而不是和平——看成是冒了“推断谬误”的风险，它对研究发现的效度带来“破坏性的后

果”。格迪斯(Geddes, 1991;也可参见 Geddes, 2003)指出,基于因变量来选择个案是一种“禁忌”,它“蛊惑着”很多重要的定性研究,包括斯考切波对社会革命的分析。有关基于因变量的个案选择研究在金、基欧汉和维巴对此问题的讨论处达到高潮,这就是本章开篇所引用的导语。

这种批评的统计学基础已经在文献中有良好的阐述,尤其是赫克曼的文章(Heckman, 1976; 1979)。在双变量关系中,与使用因变量的全部变异相比,选择在因变量上被删节的一个样本会产生一个平稳的斜率。因此,用删节过的样本来估计回归线会导致系统的误差(即偏差)。这种说法的寓意在于,研究者选择的个案样本不应该都处在因变量的一个有限范围内。在定量传承中,此建议无须争议,并且如果处理的是典型的线性回归模型的话,这一点就更容易理解。

如果研究者所用的因果模型更具有定性研究的特征,那么这个建议就不太恰当了。例如,考虑如下因果模型: $Y=(A*B*C)+(A*C*D*E)$。在这个等式中,A 是 Y 的一个假定的必要条件。如果有人想要验证关于 A 的这个假设,那么哪种个案选择策略是合适的呢?

为了回答这个问题,回想一下第 2 章给出的关于必要原因的定义,如下所示:

$$P(X=1 \mid Y=1)=1 \tag{14.1}$$

注意该等式要说明的是:考虑 $Y=1$ 的所有个案,并分析它们是否都满足 $X=1$。如果满足,则此证据就会支持必要条件假设。当然,这意味着是在因变量基础上的选择;它意味着选择这些个案恰恰是因为它们都满足 $Y=1$。

J.S. 穆勒(Mill, 1843, 1974)的求同法是思考这种设计的一个传统方式,该方法作为一种策略可用来考察关于必要条件的假设。最近,方法论者又探索了研究者必须考察多少个相符合的个案(即 $Y=1$ 的个案也满足 $X=1$)才能确信假设是有效的(Dion, 1998; Braumoeller and Goertz, 2000; Ragin, 2000)。从我们此处的目的角度讲,关键在于,当在定性研究中检验必要条件的时候,基于因变量来选择个案要直

接遵循关于必要条件的定义。①

为了分析定性的个案选择规范和统计传承之间的冲突，可以考虑格迪斯(Geddes，1991，2003)关于选择偏误(selection bias)的文章。格迪斯考察关于东亚地区(如韩国、中国台湾和新加坡)持续快速经济增长方面的定性文献中的问题。她观察到，当分析者在解释高度增长时，经常强调政府在创造训练有素且沉默的劳动力(即“劳动力压榨”)方面的角色。然而，这些学者在做出这种因果推断的时候，利用的证据只来自保持高度增长的一些国家。她建议，由于这些学者只选择极端的个案(即高度增长的个案)，他们的发现会受限于选择偏误。她提出，如果在选择个案时不考虑其在因变量上的值，那么在劳动力压榨和经济增长之间可能出现不同的关系。

为了从经验上指出此点，她用发展中国家的一个大样本来展现人均 GDP 和劳动力压榨之间的双变量数据。她提供了 1970 年人均 GDP 高于韩国的 32 个发展中国家的数据，图 14.1 再现了该数据

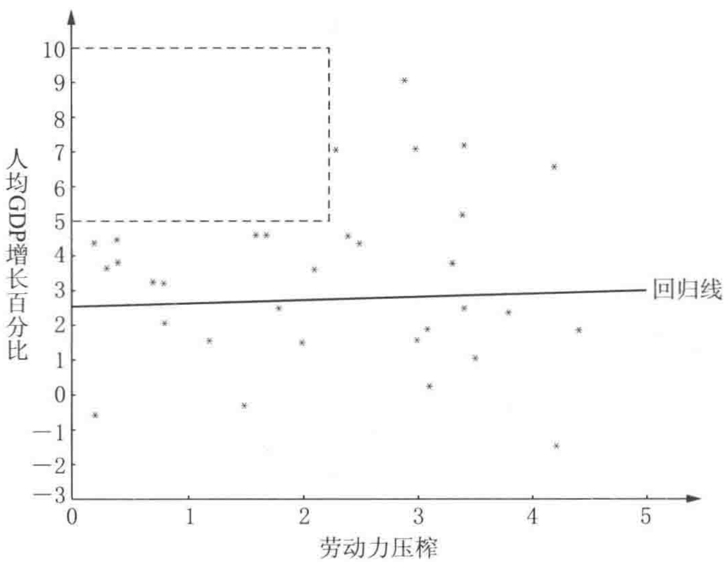


图 14.1　基于因变量的选择:劳动力压榨和经济增长

① $Y=0$ 的个案会在评价必要条件的假设上发挥作用，虽然这一点超出了本讨论的范围。特别是，它们帮助学者从不重要的必要条件中区分出重要的必要条件(参见 Braumoeller and Goertz，2000；Goetrz，2006；Ragin，2008)。

(Geddes, 2003:104)。正如图中的平缓回归线所示,这些数据并没有线性关系(斜率是 0.09, $R^2=0.003$)。因此,当格迪斯看到图 14.1 的散点图后,她认为劳动力压榨与经济增长无关。①

然而,从集合论的观点来看,该数据的确表明在劳动压制和经济增长之间可能有重要关系。数据有一个空白区域,在图 14.1 的左上半区域没有观测值,这个事实值得我们重点关注。如果下面的假设是正确的,那么空白区域就是研究者期望要看到的:劳动力压榨是经济高速增长的必要条件。因此,虽然数据不支持线性关系假设,但是它们的确与必要条件假设相一致。所有的经济异常增长的个案都至少有中等程度的劳动力压榨(墨西哥这个个案有较低水平的劳动力压榨,但有高度增长)。因此,在这个总体中,至少中等程度的劳动力压榨对于持续高度增长来说是必要条件。头脑中武装了集合论因果模型的定性研究者不能不注意到数据的这个方面。

有趣的是,如果回过来看提出劳动力压榨是持续高度增长的原因的最初学者,会发现他们趋于从必要条件而不是一种线性协变模式的角度来思考关系。例如,戴约(Deyo, 1987:182)写道,"训练有素、低成本的劳动力……已成为发展的先决条件"。库(Koo, 1987:174)断言道,"对工业劳动力的控制和规训……是边缘国家必须提供的条件,这些条件可使外国资本的投资环境优化,同时促进国内资本的商业自信"。这些引文使用了经典的必要条件语言(例如,"先决条件"、"必须提供")。因此,格迪斯的数据似乎与新兴工业国家的学者提出的实际假设相一致。

总之,个案选择是一个依赖于模型的问题;为了检验所研究的假设的经验意义,研究者必须选择个案。如果模型提出了一个必要条件,一个好的策略就是选择 $Y=1$ 的个案。如果假设涉及一个充分条件,最好的个案通常是 $X=1$ 的那些。如果假设预设有一种可加的线性模型,那么这些个案选择的模式就是不恰当的了。

① 这里直接有一个担心,即格迪斯的例子表明选择偏误导致定性研究者在他们的被删节的样本中高估了真实的关系强度。然而,就双变量来说,选择偏误应该在删节的样本中产生一种弱的关系。

14.3 选择个案的策略

在本节中，我们将用表 14.1 作为一个程式化的例子来说明观点。该表中包含两个二分变量，原因变量 X 和结果变量 Y。(0, 1)单元中没有原因但有结果，可视之为必要条件单元，因为存在一个必要条件，该单元必须是空的。(1, 0)单元是一个充分条件单元，其中有原因但没有结果，因为如果充分条件存在，该单元必然为空。(1, 1)单元包含的个案都既有原因，也有结果；(0, 0)单元的个案都缺乏原因和结果。

表 14.1 用于个案研究选择的单元

	$X=0$	$X=1$
$Y=1$	(0, 1)	(1, 1)
$Y=0$	(0, 0)	(1, 0)

14.3.1 随机选择

在定量传承中，标准的个案选择策略是随机选择。如果研究者要选择大量的个案，这种选择策略就有很多优点。最基本的优点是，随机选择确保了个案选择过程与因果模型中的任何变量都无关。用表 14.1 的术语讲，大量个案的随机选择确保了没有任何单元被给予优先权。当检验一个统计模型时，这种数据恰恰是人们想要的。

然而，当研究者进行小样本的个案研究时，随机选择的效用就有争议了。费伦和莱廷(Fearon and Laitin, 2008)主张，当评估回归结果的因果机制时，可对小样本的个案研究实施随机选择。他们认为，非随机选择的个案研究很少令人信服，因为我们不能确信，研究者是否“挑选”(cherry-picking)了那些预先了解的个案来支持假设。另外，仅仅因为人们拥有关于这些个案的好数据或知识而选择它们，这种做法引出了

代表性问题。他们相信，这种代表性问题也适用于研究者在选择观测项时所采取的策略，因为这些观测项在自变量和/或因变量上有值。为了回应这些问题，并进行更客观的检验，他们敦促个案研究者使用随机数生成器来选择他们的个案研究。①

不出意外，定性学者几乎从不使用随机选择策略。②这些研究者会部分基于个案在特殊变量上的值而有目的地选择个案。他们之所以这样做，一个重要原因在于，一定类型的个案会比其他个案在检验他们的因果模型方面提供更大的影响力。

现在，想象一位要检验关于必要原因的假设的学者。此时，图 14.1 的(0, 1)单元就非常重要了。该单元中的任何个案都会被证明是不成立的。毫无疑问，研究者会审视这些个案，分析它们是否真的挑战了假设。他/她会探索，这里是否存在着测量误差或者是否有未预料到的因果机制。

另外一个重要的单元是(1, 1)单元。研究者期望从其中的个案中看到将必要条件与结果连接起来的因果机制。定性研究者会详细探究(1, 1)个案，以便发现它们是否符合预期。在定性分析中，研究的主要目的是解释 $Y=1$ 的个案，因此，要密切关注那些也拥有一个可能为真的必要原因的 $Y=1$ 的个案，这是研究者的完全自然的想法。

在(1, 0)单元里存在着必要条件，但不存在结果，该单元的用处不大。但它还是能帮助研究者探索感兴趣的因果机制。研究者利用这些个案，可能期望必要条件显示出其因果功用来，即通过移除障碍或创造机会来促成结果或使结果成为可能。尽管由必要原因呈现出来的这些可能性在结果中不会达到顶点，但是检查它们是否会呈现仍然是合适的。困难在于，必要条件的因果机制可能难以观测到，因为其他重要条件的缺乏会阻碍结果实际发生。

① 在评估不同的个案选择策略时，赫伦和奎因(Herron and Quinn, 2011)发现，相对于立意个案选择策略来说，随机选择的效果更好。

② 正如金、基欧汉和维巴所言："定性研究者经常(适当地)改变随机选择的概念，拒绝错过重要个案的风险，这些个案可能因随机选择而不被选中。"(King, Keohane and Verba, 1994:125)

最后，定性研究者关于(0，0)单元的观点是十分矛盾的。他们因这一单元而分成两个不同的方向。一个方向是，选择 $X=0$ 可以是检验必要条件假设的一个好方式。这样的一个假设可以对这些个案做出清晰的预测：Y 不会发生。另一个方向是，在这些个案中，大多数都与检验关于必要条件的假设关系不大(Braumoeller and Goertz，2002；Clarke，2002；也见 Seawright，2002)。例如，在随机选择规则下，可以证明(0，0)单元是尤其有问题的。在定性研究中，这些单元经常充满着个案，因此，随机选择的话会选出(0，0)单元中的许多个案。在这个单元中有很多个案，因为定性研究者通常研究少见的结果，因而使得 $Y=0$ 的个案数远大于 $Y=1$ 的个案数。在有问题的(0，0)单元中，个案总量将取决于必要原因 X 的分布，但是，通常(0，0)单元是最受欢迎的，因为通常情况下，学者会研究那些不常呈现的必要原因。①

与定性传承的实践相比，赫伦和奎因(Herron and Quinn，2011)发现，因其极值或不寻常值而选个案，这样的个案选择策略对于评价统计假设是最没有帮助的。他们使用如表 14.1 所示的 2×2 表总结道，从总体稀少的单元中选择个案的策略是需要避免的。“从纯统计抽样的视角看，关注那些不代表作为一个整体的总体的个案是资源的巨大浪费。虽然这样的个案可能对于探索性分析和/或理论建构是有用的，但是，根据定义，它们可提供的关于总体层次的平均因果效应的信息量是有限的”(Herron and Quinn，2011:13)。然而，正如我们所见，在评价集合论的假设时，最有用的单元总体通常是稀疏的。例如，在必要条件假设下，很有用的(0，1)单元并不“代表作为一个整体的总体”，因为必要条件假设说的是该单元应该为空。并且，通常有最多观测值的单元，即(0，0)单元是最无用的。

在此，我们的观点是，个案选择依赖于因果模型和研究目标。一个观点是，我们研究的是平均处理效应，赫伦和奎因(Herron and Quinn，

① 当定性研究者研究充分条件时，提供最大影响力的那些单元稍有不同。对于这些假设来说，(1，0)单元现在就非常重要了，因为这一单元中的任何观测值都会否证该假说。证明(1，1)单元中的观测值仍然很重要。在(0，0)单元中的个案基本上与检验关于充分条件的假设无关，因而通常也不被选来进行个案研究分析。

2011)围绕这个观点明确地建构他们的分析。因此,他们的结论和建议不能与那个特定的研究目标相分离。同样,在定性研究中,我们关于个案选择实践的讨论也假定我们处理的是集合论假设。定性的选择实践只有在那样的情境中才有意义。

14.3.2　实质重要的个案

另一种考量涉及研究者应该选择的一些特定个案,即一旦抉择,一组个案就是有用的。例如,在(1, 1)的个案中,定性研究者应该选择哪些个案呢?按照费伦和莱廷(Fearon and Laitin, 2008)的观点,人们可能建议在这些个案中随机选择。该建议再一次建立在来自大样本(large-N)分析的观念之上,其中不存在事先(ex ante)重要的个案。随机选择帮助研究者避免有意或无意地挑选那些支持有利结果的个案。

然而,定性研究者即使在来自一个有用单元(例如,(1, 1))的个案之间,也从来不用随机选择。相反,他们经常选择自己非常熟知或很容易熟知的个案。在定性传承中,熟知一个个案可以显著地增进个案内分析(参见第二部分),提升人们做出有效推断的机会。

定性研究者相信,除了考虑个案的专业知识的价值之外,有些个案比其他个案更有"实质重要性"。如果使用随机选择,选择到这些实质重要的个案的几率通常较低。实质重要的个案经常是一种现象的"理想型"或有名的示例。同样,具有特殊的规范旨趣或具有重要政治角色的个案可以被认为有实质重要性。

好的定性理论必须能够解释实质重要的个案。不能解释时,对它们不利。毕竟,如果研究的目的是准确地解释个案,那么一个理论在解释重要个案上的能力应该得到高度评价。只能解释较少个案的理论则较少受认可。格洛德斯通(Glodstone, 2003:45—46)提供了一个很好的例子,他讨论了马克思主义理论不能充分解释法国大革命的后果:"马克思主义观点可能在其他个案中仍然成立,但是当发现它不能解释历史上最重大的一次革命(即在当时最大、最有影响力和最有效仿性的国家中发生的这次革命,并且它常常是马克思主义理论的范本)时,当

然会动摇人们对该理论的价值的信念。”

从统计的视角来看，相比之下，规范的做法是当检验假设时所有的个案都应该被先验地赋予同等权重。由于理论的目标不是解释特殊个案，而是将因果效应向大样本推广，因此，就没有理由特别关注特定的个案。当评价一个变量在社会革命上的效应时，法国大革命不应该算做例外。如果很多其他个案是符合的，那么法国大革命的不符合就不是一个特殊问题，或者至少不比其他革命——如玻利维亚革命（Bolivian revolution）——更成问题。因此，从定量传承的视角看，定性学者对实质重要个案的担心似乎令人困惑。

14.3.3 使用跨个案证据来选择个案研究

在多元方法研究中，挑选个案研究的时候，标准的策略涉及使用跨个案的证据来区分尤其有用的个案。大多数此类文献都假定研究者首先进行大样本的统计分析，然后用其结果来分出个案研究（例如，Lieberman，2005；Seawright and Gerring，2007，2008）。出于同样的原因，挑选出个案研究作为一个随机样本：在感兴趣的维度上要有代表性和变异性。当研究者努力运用个案研究的时候，通常要假定这有助于他向一个得到良好界定的大的个案总体做推广。在这一情境下，个案研究既可精炼又能检验统计模型。

在多元方法方面的文献中，个案选择策略菜单中有多种选项（例如，Seawright and Gerring，2007 提出了 9 种）。这里的假定在于，对个案研究的分析将涉及定性研究。然而，也有假定指出，统计模型中变量效应的估计值是分析的起点和结点。因此，个案选择的过程完全依赖于统计模型。例如，选择一个“典型”或“在线”个案的策略涉及发现一个低残差（low-residual）的观测，此观测项在因变量上的值或多或少匹配着它的回归预测值。“异常”或“离线”的个案则是残差高的个案，它在因变量上的值与回归预测值不接近。我们相信，这些类型的个案研究作为回归研究的补充是十分必要的，并且多元方法方面的文献很好地阐释了其原因（Collier，Brady and Seawright，2010）。

对于探索集合论假设的定性研究来说，哪些定量的多元方法个案选择策略最有用？一方面，可以简单地认为这些策略与这种研究无关。在这些技术中，大多数都与关于必要条件和“充分非必要条件的必要非充分部分”(INUS conditions)的假设很少相干。此外，某些定性研究甚至无意于向大总体推广。在已知它们的范围有限的情况下(参见第16章)，定性研究者有时候在大多数或全部相关的总体中进行个案研究。因此，从大总体中选择个案研究，这种想法恰恰可能不适用。

然而，另一方面，有些技术可以延伸到定性研究中来评价集合论假设。一条基本规则是使用在线或离线的个案，这条规则修正后可应用于集合论的情境中(参见 Schneider and Rohlfing，2010)。例如，我们可以回到图14.1中的格迪斯数据。其中的假设指出，至少适当的劳动力压榨对于高度增长来说是必要的。就这个假设而言，所有的个案都“一致”，可能墨西哥是个例外。然而，在这些一致的个案中，并不是全部都应该看做在线的例子。真正在线上的个案是那些有原因和结果的强实例。因此，位于右上区域的个案才是“典型个案”的最佳候选者。在格迪斯的数据中，巴西、马来西亚、新加坡、韩国、叙利亚和中国台湾是6个高劳动力压榨和高增长的国家和地区，因而是最佳的在线个案。除了叙利亚以外，在关于新兴工业化国家(newly industrializing countries，缩写为 NICs)的文献中，它们也是常常被选择的个案。①

在西赖特和耶林(Seawright and Gerring，2007)讨论的策略中，可能有两个策略最容易应用在定性研究中，它们是“关键个案”(critical case)策略和“途径个案”(pathway case)策略(也可参见 Gerring，2007)。②关键个案被定义为是那些“最可能或最不可能展示一个既定结果”的个案(Seawright and Gerring，2007：89)。在讨论图14.1的时候，我们指出有一个特殊单元对于关于必要或充分条件的假设有重要后果。这个单元中的个案就是关键个案，因为它们最不可能展示出所

① 相比之下，离线或偏离的个案则位于散点图的左上象限(在格迪斯的数据中，不存在主要的离线个案)。

② 西赖特和耶林在后续研究中设定的明确目标是向大样本总体做推广，其中放弃了这两种策略。参见西赖特和耶林(Seawright and Gerring，2008)。

关注的结果(即这些个案缺乏一个必要条件)或最可能展示出结果(即个案具有一个充分条件)。[①]

相比之下，途径个案被定义为是那些"X_1 而非 X_2 可能已经产生一个正向结果($Y=1$)"的个案。这些个案研究的目的是探究因果机制而不是检验一般的命题。我们相信，这类个案研究在定性研究中很普遍。个案研究的中心目标通常是评价备择的论点，需要利用过程追踪法来探究因果机制。就图 14.1 而言，通常在(1，1)单元中可以发现途径个案。研究者分析的个案要有(多个)原因或所关注的结果，其目标是决定因果过程在个体观测值内部是否如预期那样起作用。由于这类个案研究非常普遍，因此，它本身在第 9 章中得到了实质性的关注。

14.4 结论

本章的讨论表明，个案选择的规则要服从研究目标。在定性和定量研究之间，研究目标的不同导致了不同的关于最佳个案选择策略的观点。

首先，在定量传承中，随机选择受到高度欢迎，并且是选择大样本个案完全适用的模式。这一选择策略在 2×2 表中没有特别偏好的单元，并且可以避免统计分析中的偏误。然而，如果检验的是一个集合论假设，那么特定种类的个案要比其他个案更加重要；并不是所有 2×2 表中的单元都有同等重要性。因此，在定性研究中，研究者通常对某些单元而非其他单元的抽样有很强的偏好。

其次，一旦认定一组个案在分析上是有用的，定性研究者便经常诉诸那样的个案，他们有关于个案的专业知识并且能最有效地执行个案

① 尽管关键个案策略在定性研究中很常见，但它可能不总是统计分析的好的补充。在赫伦和奎因(Herron and Quinn，2011)评价西赖特和耶林的策略时，他们对一种版本的关键个案策略持有消极观点，认为它是"无成功希望"的，因为这种策略要求先对结果有很深的认识。

内分析。除此之外，他们经常极为关注实质重要的个案。这与定量研究有很大差异，后者并不会基于规范的或历史的身份而赋予特定的个案以特殊的权重。这种差异根植于两种传统的对比性的研究目标：定性研究者旨在解释特殊个案，因此十分在意某些特定的个案；定量研究者则试图估计大总体内的平均效应，因而不会在意特定的个案本身。

最后，当评价一个关于平均效应的大样本（large-N）假设时，有关选择个案的定量多元方法文献提供了一系列有用的策略。然而，其中的很多策略不能没有问题地推广到定性研究中来检验集合论关系。它对于选择“典型的”或“偏离的”个案来说意味着什么，取决于正待检验的因果模型。在统计模型中，这些类型个案的定义要考虑它们的残差。但是在一个集合论假设中，典型的个案将是同时呈现原因（或者原因序列）和结果的观测项。偏离的个案（就必要条件来说）是有原因但无结果的观测项。在此方面，“典型”和“偏移”的意义恰恰依赖于研究者是在使用统计模型还是集合论模型。

参考文献与推荐读物

在统计分析中，关于选择偏误的经典读物参见赫克曼（Heckman，1976，1979）；马达拉（Maddala，1983）；阿肯（Achen，1986）和曼斯基（Manski，1995）。一种观点认为，统计分析中的选择偏误并不像人们声称的那样是一个严重的问题，参见斯托曾伯格和赖莱斯（Stolzenberg and Relles，1990）。有关研究者用选择偏误来批判定性研究的例子，参见阿肯和斯尼达尔（Achen and Snidal，1989）；格迪斯（Geddes，1991，2003）；金、基欧汉和维巴（King，Keohane and Verba，1994：第4章）。对这些批判的反驳，可以参见科利尔和马奥尼（Collier and Mahoney，1996），科利尔、西赖特和马奥尼（Collier，Seawright and Mahoney，2004）。

有关随机选择在大样本研究中的优势，简要的讨论可以在金、基欧汉和维巴（King，Keohane and Verba，1994：124）中找到。有关倡导随机选择小数量个案的讨论，可以参见费伦和莱廷（Fearon and Laitin，2008），赫伦和奎因（Herron and Quinn，2011）。有关个案选择的多元

方法的重要文献包括利伯曼(Lieberman，2005)；西赖特和耶林(Seawright and Gerring，2007，2008)；罗尔夫林(Rohlfing，2008)和塞科恩(Sekhon，2004)。在检验集合论假设的个案选择方面，相关的讨论可见布劳米勒和格尔茨(Braumoeller and Goertz，2002)；克拉克(Clarke，2002)；弗里德曼(Freedman，2010)；马奥尼和格尔茨(Mahoney and Goetrz，2004)；拉金(Ragin，2000)；施奈德和罗尔夫林(Schneider and Rohlfing，2010)和西赖特(Seawright，2002)。

Achen，C. H. 1986. *The Statistical Analysis of Quasi-experiments*. Berkeley：University of California Press.

Achen，C.H.，and D. Snidal. 1989. "Rational Deterrence Theory and Comparative Case Studies." *World Politics* 41：143—169.

Braumoeller，B.，and G. Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44：844—858.

Braumoeller，B.，and G.Goertz. 2002. "Watching your Posterior：Comment on Seawright." *Political Analysis* 10：198—203.

Clarke，K.A. 2002. "The Reverend and the Ravens：Comment on Seawright." *Political Analysis* 10：194—197.

Collier，D.，H.E. Brady，and J. Seawright. 2010. "Introduction to the Second Edition：a Sea Change in Political Methodology." In *Rethinking Social Inquiry：Diverse Tools，Shared Standards*，2nd edition，edited by H.E.Brady and D.Collier. Lanham，MD：Rowman & Littlefield.

Collier，D.，and J. Mahoney. 1996. "Insights and Pitfalls：Selection Bias in Qualitative Research." *World Politics* 49：56—91.

Collier，D.，J. Seawright，and J. Mahoney. 2004. "Claiming too much：Warnings about Selection Bias." In *Rethinking Social Inquiry：Diverse Tools，Shared Standards*，edited by H. E. Brady and D.Collier. Lanham，MD：Rowman & Littlefield.

Deyo，F.C. 1987. "State and Labor：Modes of Political Exclusion

in East Asian Development." In *The Political Economy of the New Asian Industrialism*, edited by F.C.Deyo. Ithaca: Cornell University Press.

Dion, D. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30:127—145.

Eckstein, H. 1975. "Case Study and Theory in Political Science." In *Handbook of Political Science*, vol.7, edited by F.I.Greenstein and N.W.Polsby. Reading, MA: Addison-Wesley.

Fearon, J.D., and D.D. Laitin. 2008. "Integrating Qualitative and Quantitative Methods." In *Oxford Handbook of Political Methodology*, edited by J.M.Box-Steffensmeier, H.E.Brady, and D.Collier. Oxford: Oxford University Press.

Freedman, D.A. 2010. "Black Ravens, White Shoes, and Case Selection." In *Statistical Models and Causal Inference: A Dialogue With the Social Sciences*, by David Freedman and edited by D.Collier, J.Sekhon, and P.Stark. Cambridge: Cambridge University Press.

Geddes, B. 1991. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." In *Political Analysis*, vol.2, edited by J.A.Stimson. Ann Arbor: University of Michigan Press.

Geddes, B. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.

George, A.L., and R. Smoke. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.

Gerring, J. 2007. "Is there a (viable) Crucial-case Method?" *Comparative Political Studies* 40:231—253.

Goertz, G. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social

Science." *Studies in Comparative International Development* 41: 88—109.

Goldstone, J.A. 2003. "Comparative Historical Analysis and Knowledge Accumulation in the Study of Revolutions." In *Comparative Historical Analysis in the Social Sciences*, edited by J.Mahoney and D.Rueschemeyer. Cambridge: Cambridge University Press.

Heckman, J.J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for such Models." *Annals of Economic and Social Measurement* 5:475—492.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153—161.

Herron, M.C., and K.M. Quinn. 2011. "A Careful Look at Modern Qualitative Case Selection Methods." Paper presented at the annual meeting of the Midwest Political Science Association, Palmer House Hilton, Chicago, IL.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Koo, H. 1987. "The interplay of state, social class, and world system in East Asian development: the cases of South Korea and Taiwan." In *The Political Economy of the New Asian Industrialism*, edited by F.C.Deyo. Ithaca: Cornell University Press.

Lieberman, E. 2005. "Nested Analysis as a Mixed-method Strategy for Comparative Research." *American Political Science Review* 99:435—452.

Maddala, G.S. 1983. *Limited-dependent and Qualitative Variables in Economics*. Cambridge: Cambridge University Press.

Mahoney, J., and G. Goertz. 2004. "The Possibility Principle: Choosing Negative Cases in Comparative Research." *American Politi-*

cal Science Review 98:653—669.

Manski, C.F. 1995. *Identification Problems in Social Sciences*. Cambridge: Harvard University Press.

Mill, J.S. 1843, 1974. "Of the Four Methods of Experimental Inquiry." In *A System of Logic*, book 3, chap. 8. Toronto: University of Toronto Press.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Rohlfing, I. 2008. "What You See and What You Get: Pitfalls and Principles of Nested Analysis in Comparative Research." *Comparative Political Studies* 41:1492—1514.

Schneider, C.Q., and I. Rohlfing. 2010. "It's Complex! Combining Set-theoretic Methods and Case Studies in Multimethod Research." Manuscript.

Seawright, J. 2002. "Testing for Necessary and/or Sufficient Causation: Which Cases are Relevant?" *Political Analysis* 10: 178—207.

Seawright, J., and J. Gerring. 2007. "Techniques for Choosing Cases." In *Case Study Research: Principles and Practices*, by J.Gerring. Cambridge: Cambridge University Press.

Seawright, J., and J. Gerring. 2008. "Case-selection Techniques in Case Study Research: a Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61:294—308.

Sekhon, J. 2004. "Quality Meets Quantity: Case Studies, Conditional Probability, and Counterfactuals." *Perspectives on Politics* 2: 281—293.

Skocpol, T. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge

University Press.

Stolzenberg, R.M., and D.A.Relles. 1990. "Theory Testing in a World Constrained by Research Design: the Significance of Heckman's Censored Sampling Bias Correction for Nonexperimental Research." *Sociological Methods and Research* 18:395—415.

15 概括

概括：1. a. 归纳为一般的形式、类别或规律。b. 排除不确定性或不具体性。2. a. 从殊相中推断。b. 从……中得出推断或一般性的结论。3. a. 使……总体上或普遍地可应用。b. 使……普及。

——《美国传统英语词典》

覆盖律有一个令人尴尬的不足；在如生物学、心理学等科学和社会科学中，几乎没有什么可观测的规律被认为是解释性的。

——彼得·赫德斯特伦(Peter Hedström)

彼得里·于利科斯基(Petri Ylikoski)

15.1 导言

大多数社会科学家都想对世界做出概括。研究任务的核心目标是生产能够超越时空的概念、模型和理论。然而，不管是哲学文献，还是社会科学方法论文献，都对概括可以采取的特定形式所言甚少。在本章中，我们比较在定性和定量研究中使用的典型概括模式。

第一个问题是追问，人们所说的“概括”是什么意思？很多人会认识到描述性概括与因果性概括之间的差异（如 King, Keohane and Verba, 1994）。描述性概括经常涉及一个变量，它在一个总体中“描

述”事件的某种状态。例如，研究者可以做一项调查，并用其发现得出概括性结论：在特定的一天，50%的美国人对奥巴马总统有好感。相比之下，因果性概括总是涉及至少两个变量，A 和 B。理想条件下，因果性概括会指出在一个总体中 A 和 B 的关系形式和强度。

在定性和定量传承中，研究者做出的因果概括可以假定采取不同的形式，至于什么才算做强概括，他们也采纳了对比性的理解。定量研究者使用关联性和平均处理效应等观念来思考概括。相应地，概括的优势与关联强度和处理效应的规模有关。

相比之下，在定性研究中，因果性概括经常假定采用一个集合论形式：“所有/没有 A 是 B。”当然，有时候有一些反例，因此研究者会使用如“差不多”或“几乎”全部或者没有 A 是 B 之类的语言。例如，有人可能注意到，在民主政体的国家从来不发生种族灭绝事件。这会导致这样的概括：“所有的种族灭绝事件都是非民主国家的事件。”

因此，在每一种传承中，有力的(因果)概括意味着不同之事。在定量传承中，它意味着变量之间有强的相关或处理效应。在定性传承中，它意味着一个类别中的成员资格对于另一类别中的成员资格来说是必要的，或者说是保证性的。

两种传承相互很难看到和分析对方的典型类型的概括。定性方法的设计不是为了发现相关、关联或平均处理效应。就此而言，标准的统计方法的设计也不是为了研究集合论的概括。在本章中，我们会指出，在相当多的文献中存在着几乎完美的集合论概括的例子。尽管这一发现可能使某些吃惊，但它在定性研究中是可以理解的。

15.2 定性的概括

在定性研究中，集合论关系与必要及充分条件有着密切的亲和性。在集合论语言中，“所有的 A 都是 B”意味着 A 是 B 的子集。用逻辑学术语来陈述，即 B 是 A 的必要条件。有趣的是，这些概括有着基本的

覆盖律形式。例如,蕴含在民主式和平中的因果性概括很容易被转化成律则式的覆盖律(canonical covering law)形式(例如,Hempel, 1942; Hitchcock and Woodward, 2003):

民主国家之间不发生战争。
美国和加拿大是两个民主国家。
因此,
在美国和加拿大之间不会发生战争。

在科学哲学中,规律通常采用集合论的概括形式。例如,阿姆斯特朗(Armstrong, 1983)关于"自然律"(laws of nature)的著名分析最关注如下观点,即覆盖律意味着所有的 *F* 都是 *G*。准确地说,"所有/没有 *A* 是 *B*"这样的强概括是覆盖律的潜在候选项。

毫无疑问,一些读者关注的是,像"所有/没有 *A* 是 *B*"这样强的概括在社会科学中很少或不存在(尽管在自然科学中可能存在)。然而,我们在研究和阅读过程中,时常发现这些集合论的经验概括的例子。下面的列表反映了我们感兴趣的主题,如国际冲突、内战、民主和经济发展。为了强调这些概括中的"全部"或"没有"用语,我们在相关的语词上加黑显示。

普遍选举权的引入在几乎所有地方(美国除外)都导致了社会主义党的发展。(Duverger, 1954:66)

民主国家中没有饥荒。(我们改版的 Dreze and Sen, 1989 观点)

富裕的民主制(几乎)不会转变为威权制。(我们对 Przeworski et al., 2000 的著名发现的表述)

自第一次世界大战以来,**没有例证**表明*拥有弱制度的不完全民主化国家*会参与,更不用说发动一场外部战争……从 1945 年开始,在 63 个经历了不完全民主转化的弱制度国家中,**没有一个国家**发动或者参与过外部战争。(Narang and Nelson, 2009:363,重

点字为这里所加）

最后的概括是就态度的系列变化给出一个陈述，这些态度发生在我所知道的开始使用大麻寻求快感的每一个个案中。(Becker, 1953:236)

我们表明，至少在拉丁美洲，没有哪一个国家的民主是因为选择使用(人权)试验而受到破坏的。(Sikkink and Walling, 2007:442)

有组织的工人阶级几乎在任何地方的民主发展中都是一个重要行动者，仅有的例外是存在于一些小国家中的少数农业民主。(Rueschemeyer, Stephen and Stephens, 1992:270)

经济危机伴随着这里述评的每一次转型，这当然不是巧合。这一模式说明，经济危机对于威权政体的破裂来说可能是一个必要尽管不充分的刺激。(Bermeo, 1990:366)

的确，自第二次世界大战结束以来，总统的党派性对收入不平等显然有一致的影响。在此期间，在六位共和党总统中每一位的治理下，80/20 的收入比提高了。相比之下，在五位民主党总统中，有四位总统(除了吉米·卡特以外)的主政带来收入不平等的降低。(Bartels, 2008，引自 Tomasky, 2008:45)

第一，只有国际化的坚定反对者才在东亚地区追求核武器，……第二，在中东地区，所有的核项目都被领导者所启动，这些领导人操纵着进口-替代和相对封闭的政治经济。(Solingen, 2008:18)

在过去的二十年里，发起自杀式运动的每个小组都有一个主要目标，……迫使在被恐怖分子看做自己祖国之处驻军的外国势力撤离。(Pape, 2005:21)

在 2001 年到 2005 年间发生的 38 起活跃的武装冲突中，仅有一起发生在世界上 1/4 的富裕国家中；基地组织在 2001 年 9 月 11 日袭击了美国。(Buhaung and Gleditsh, 2008:218)

其中的一些概括是相当著名的；实际上，在社会科学中，很多最重要的

发现都是集合论的概括。如果我们很留意这些概括的形式,不难发现更多。

在定性方法论中,穆勒的求同法(methods of agreement)和求异法(methods of difference)是研究个体变量之间关系的集合论工具。当表述因果概括的时候,自然要考察所有 $Y=1$ 的个案,观测当 Y 出现时是否有某个普遍的因素 X 也总是出现。对于求同法来说,这是标准的方案。用我们的术语来说,求同法被用来寻求这样的 X,即使得"所有的 $Y=1$ 都是 $X=1$"。贝克尔(Becker)的名篇就是这方面的一个经典例子。他的注意力是关注一些为求快感而吸食大麻的人,发现了他们中间普遍具有一系列行为。

在其中的一些例子中,作者在解释集合论概括时,认为它反映了一种因果关系。例如,贝尔梅奥(Bermeo)认为,在危机和威权政体的崩溃之间存在一种因果关系。她利用集合论的经验概括支持这一观点。显然,想要确证这种因果解释,她就需要进行额外的分析,如个案内分析和过程追踪。但是,概括本身就支持了一个变量引起另一变量的观点。

最后需要强调的是,进行这些概括的学者都假定它们适用于某些情境而非其他情境。的确,定性概括的范围经常是相当受限的(参见第16章)。此外,很显然其中的一些概括都有例外。因此,当学者提出一个遵循集合论形式的概括时,他们不认为概括就像宇宙的一个规律那样不能被违反。

15.3 集合论概括和 2×2 表

由于集合论的(因果)概括包括两个变量,因而很容易用 2×2 表来呈现这种概括(出于本章的目的,假定变量或概念是二分的)。表 15.1 利用了我们在下一节继续使用的一个例子;它是戈曼斯(Goemans, 2000)的一项研究,探讨了国际危机和战争结束后领导人的命运。一个

因变量是领导者在战争结束后是否受罚(即被流放、监禁或杀掉)。关注的一个关键自变量为外国势力是否推翻了原有的政权。表 15.1 汇总了我们的双变量发现。

表 15.1　集合论的概括:强加的政权更迭后领导人的命运

	国内强加的更迭	外国强加的更迭
受惩罚	27	22
不受罚	182	0

资料来源:Goemans，2000。

这个表表明,有一种简单的方法可以发掘强的定性概括,即在 2×2 或 $N\times N$ 表中寻找一个空单元(或者几乎为空的单元)。如果存在这样的一个单元,人们就用"所有/没有 A 是 B"这种形式来表示该表格的核心关系。在戈曼斯的分析例子中,我们就可以得出结论,即"所有外国强加的政权更迭都导致前领导人受罚"。

在第 2 章中,我们比较详尽地讨论了 2×2 表集合论分析对比统计分析,因此,我们在这里仅总结两个关键点。

1. 当面临一个集合论的关系时,不同的统计关联测度将因它们解释关系的方式而变。一些常见的 2×2 关联测度(如 χ^2，τ_β)会将这种关系看成是显著的但不是很强。相比之下,发生比率这个关联测度会表明一个非常强且显著的关系。

2. 统计测度通常不在"所有 X 都是 Y"和"所有 Y 都是 X"这两类概括之间做出区分。例如,发生比率并不区分充分条件和必要条件。

核心点在于,统计检验不能精确地报告数据的某些特点。相反,关键在于设计的这些检验并不是为了分析集合论概括。当然,由于同样的原因,集合论工具的设计也不是为了分析在定量传承中通常研究的那类对称关系。

15.4 统计模型、完美预测项和集合论概括

在统计分析中有一个众所周知的问题，即我们所谓的“完美预测项”(perfect predictors)问题。[①]基本上讲，在极大似然估计中，如果一个自变量完美地预测了结果，那么统计方程就不能被估计(参见 Zorn，2005 的详尽讨论)。在 20 世纪 90 年代，一些流行的软件包，如 SAS 第 6 版忽视了这个问题，因而当该问题出现时会产生无意义的结果。目前的统计软件包会提出警告(如 SAS)或者通过警告将干扰变量从模型中移除(如 Stata)。在 R 平台下，软件会估计这个模型，并把它留给研究者，让其通过注意异常大的标准误来发现该问题。

一旦发现该问题，大多数学者都会从模型中简单地去掉干扰变量(Stata 会自动消除)。这样的解会引发我们所谓的完美预言项悖论(Paradox of the Perfect Predictor)。

> 完美预测项悖论。通常到目前为止，拥有最强因果效应的变量会被从模型中移除。

学者不关注一个异常的强关系，而把它看做一个需要解决的统计学问题。虽然有一些方法可以处理此“问题”，[②]但是到目前为止，最常见的解是在模型中去掉干扰变量。例如，从统计学模型中去除一项的理由可做如下考虑：“ICOW(Issue correlates of War)数据的使用要求模型设定要有所改变。核心地位之间的交互项被从模型中去掉，因为其 0 值也完美地决定了因变量的 0 值。”(Gartzke and Jo，2009：224)

---

① 在统计学中(例如，Heinze and Schemper，2002)，它通常被称为“分离”问题。完美预测项将结果变量完全分为“0”组和“1”组。一般来说，当在一个给定的单元中有一个 0 值时，就会发生这种情况。

② 佐恩(Zorn，2005)讨论了几种针对这些变量获得“更合理的”参数估计值的方法。贝叶斯技术在解决这个问题时，利用先验值使估计似然方程成为可能。

表 15.2　完美预测项:外国强加的政权更迭

变　　量	MLEs		MPLEs	
	$\hat{\beta}$	发生比率 (Odds ratio)	$\hat{\beta}$	发生比率 (Odds ratio)
常　　量	−2.96 (0.46)	—	−2.87 (0.44)	—
其他小的战败国	0.85 (0.66)	2.3	0.85 (0.63)	2.3
其他大的战败国	3.36 (1.02)	28.8	3.20 (1.00)	24.5
混合政权的小战败国	2.69 (0.62)	14.8	2.61 (0.61)	13.7
混合政权的大战败国	3.24 (0.89)	25.6	3.12 (0.87)	22.5
外国强加的政权更迭	22.85 (4 840.20)	8.4×10^{9}	5.49 (1.51)	243.0

资料来源:Zorn，2005:167。

表 15.2 呈现了佐恩(Zorn，2005)对戈曼斯数据的分析。如在表 15.1 中所见，这个完美预测项就是“外国强加的政权更迭”。其参数估计值是 8.4×10^{9}，即对这样一位领导人的惩罚比率大约比在没有外国干涉情况下被免职的领导人受罚的比率多 84 亿倍。佐恩发现了无限大——从数学上讲，84 亿基本上就是无限大——是一个不合理的估计值，并且给出一些达到更小值的方法。他认为他的修正方法就是“就外国强加的政权更迭对战后领导人命运的影响，呈现一幅非常可信的图景”(Zorn，2005:167)。在更“真实”的模型中，发生比率“只有”243(参见表 15.2)。尽管 243 远远小于 84 亿，但它仍然很高，以至于这些结果对于几乎所有的实践目的来说都意义相同。①

完美预测项表明这两种传承是如何对同一数据有不同反应的。定性学者容易使完美预测项成为他或她的分析的中心。定性学者之所以

① 当标准误很大时，也可作为另一种区分出完美预测项的方式。例如，由表 15.2 可见，干扰变量的标准误超过了 4 800。

追求完善预测项，是因为其目标是全面地解释自身感兴趣的结果。相比之下，在定量领域，分析的目标很可能是估计除了完美预测项以外的其他变量的效应。有此目标后，完美预测项的存在就成了一个统计学问题，这个问题会要求剔除变量，或者引入其他的方法来解决。

15.5 控制变量和完美预测项

很多文献都论述了有观测数据的因果推断，大多数都关注混淆变量问题。加入一个新的控制变量后，一个关键的统计发现会消失，这种情况总可能发生。具备怀疑精神的评论人很快会指出作者可能没有加入一个关键的控制变量。经济学家可能着迷于这些忽略变量偏误问题，因而要针对每一年及每一个区域的截面加入控制和固定效应。正如利伯森和林恩所言：

> 影响因变量成立的条件或因素几乎有无数个(用当代社会学语言来讲)。如果一次调查产生了一种复杂的分析，其中比如要考虑 15 个变量，那么在当代的分析中一个完全可以接受的提议是：将第 16 个变量也考虑进来。“控制”了额外的属性或许会完全改变先前得到的结论，这种可能性总是存在的。(Lieberson and Lynn，2002：8)

由于统计结果对特殊的模型设定有敏感性，因此，近些年来，为社会科学期刊撰写稿件的作者们常常不惜笔墨，用大量篇幅展现“稳健性”分析。这些讨论通过改变统计分析中的诸多核心特点来分析主要变量是否保持其符号并在统计上显著，以此来处理统计发现的脆弱性，不过这些讨论常常被重要的网址所遮蔽。

相比之下，定性传承中提出的集合论概括本质上是双变量的概括，不要求加入控制变量。在关于必要性或充分性的声称中内含着这样的

假设,即其他变量或变量组合的效应都不能胜过一个必要和/或充分原因的效应(Seawright, 2002:181)。实际上,集合论概括对于虚假相关问题来说是稳健的。总之,我们不必担心额外引入的变量会消除原来的关系:

> 控制变量或混合变量不会使一个强的集合论概括失效。①

想要明白为什么这是正确的,可以回想一下,控制变量要在一个总体的子群体内部寻找差异性的关系。但是,当关注点从全总体转向子总体后,集合论概括仍保持稳定。如果对于总体 Z 来说,所有的 A 都是 B,那么对于 Z 的所有子总体来说,同样的概括依然成立。在全总体中的一个完美预测项在子总体中也是一个完美预测项。

考察表 15.2 中完美预测项的因果影响,我们也可以直觉地看出这一点。发生比率或者本质上是无穷大,或者是非常大,达到 243。虽然可以引入不同的控制变量,使用固定效应或者做其他的模型调整,但是因果效应也不大可能消失。就实际的完美预测来说,加入的控制变量将不会有影响。如果集合论概括有一些反例,估计的参数值会降低(由于有多重共线性),但是即使在这里,控制变量通常也不会有太大影响。

15.6 结论

强概括的观念在定性和定量传承中有不同的意义。就定性传承来说,它意味着一种集合论的关系,大致采取这样的形式:"所有/没有 A 是 B。"相比之下,在定量传承中,强概括观念表明在两个变量之间有强的统计相关,或者存在一种实质上和统计上显著的平均处理效应。

这些不同类的概括与两种传统的总目标紧密相关。定性研究者经

① 如果控制变量本身与概括变量完全相关,这就是该规则的例外之处。

常寻求全面地解释结果,包括鉴别出对于这些结果来说是必要的因素。这种取向相当自然地将本身指向对强集合论概括的追求。相比之下,在定量传承中,研究者力图估计一些特殊变量的平均效应。在此目标下,一个强概括经常陈述一个因果效应的规模怎样,或者关于因果效应的发现的稳定性如何。

参考文献与推荐读物

金、基欧汉和维巴(King, Keohane and Verba, 1994)从定量视角出发,很好地讨论了描述性概括和因果性概括。有关集合论概括和定性研究的关系,参见拉金(Ragin, 2000)。有关必要条件的理论概括和假设,参见格尔茨(Goertz, 2003)提供的一系列例证。戴恩(Dion, 1998),布劳米勒和格尔茨(Braumoeller and Goertz, 2000)讨论了用标准的统计工具来研究这些概括存在的挑战。珀尔(Pearl, 2000:第1章)讨论了集合论概括的例外情况如何导致学者去关注概率。

Armstrong, D.M. 1983. *What is a Law of Nature*? Cambridge: Cambridge University Press.

Ashworth, S., J. Clinton, A. Meirowitz, and K. Romsay. 2008. "Design, Inference, and the Strategic Logic of Suicide Terrorism." *American Political Science Review* 102:269—273.

Bartels, L. 2008. *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton: Princeton University Press.

Baumgartner, F., C. Breunig, C. Green-Federsen, et al. 2009. "Punctuated Equilibrium in Comparative Perspective." *American Journal of Political Science* 53:603—620.

Becker, H. 1953. "Becoming a Marihuana User." *American Journal of Sociology* 59:235—242.

Bermeo, N. 1990. "Rethinking Regime Change." *Comparative Politics* 22:359—377.

Braumoeller, B., and G. Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44:

844—858.

Buhaung, H., and C. Gleditsch. 2008. "Contagion or Confusion? Why Conflicts Cluster in Space." *International Studies Quarterly* 52: 215—233.

Chapman, T., and P. Roeder. 2007. "Partition as a Solution to Wars of Nationalism: the Importance of Institutions." *American Political Science Review* 101:677—691.

Cooper, G. 1998. "Generalization in Ecology: a Philosophical Taxonomy." *Biology and Philosophy* 13:555—586.

Dion, D. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30:127—145.

Doyle, M. 1983a. "Kant, Liberal Legacies, and Foreign Affairs, Part I." *Philosophy and Public Affairs* 12:205—235.

Doyle, M. 1983b. "Kant, Liberal Legacies, and Foreign Affairs, Part II." *Philosophy and Public Affairs* 12:323—353.

Drèze, J., and A.Sen. 1989. *Hunger and Public Action*. Oxford: Oxford University Press.

Duverger, M. 1954. *Political Parties: Their Organization and Activity in the Modern State*. London: Methuen.

Gartzke, E., and D.Jo. 2009. "Bargaining, Nuclear Proliferation, and Interstate Disputes." *Journal of Conflict Resolution* 53:209—233.

Geddes, B. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.

Goemans, H. 2000. "Fighting for Survival: the Fate of Leaders and the Duration of War." *Journal of Conflict Resolution* 44:555—579.

Goertz, G. 1994. *Contexts of International Politics*. Cambridge: Cambridge University Press.

Goertz, G. 2003. "The Substantive Importance of Necessary Condition Hypotheses." In *Necessary Conditions: Theory, Method-*

ology, and Applications, edited by G.Goertz and H.Starr. Lanham, MD: Rowman & Littlefield.

Goertz, G., T. Hak, and J. Dul. Forthcoming. "Ceilings and Floors: Theoretical and Statistical Considerations When the Goal is to Draw Boundaries of Data, not Lines Through the Middle." *Sociological Methods and Research*.

Hale, H. 2004. "Divided We Stand: Institutional Sources of Ethnofederal State Survival and Collapse." *World Politics* 56:165—193.

Hansen, H., S. Mitchell, and S. Nemeth. 2008. "International Organization Mediation of Interstate Conflicts: Moving Beyond the Global vs. Regional Dichotomy." *Journal of Conflict Resolution* 52: 295—325.

Harff, B. 2003. "No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder since 1955." *American Political Science Review* 97:57—73.

Heinze, G., and M. Schemper. 2002. "Solution to the Problem of Separation in Logistic Regression." *Statistics in Medicine* 21: 2409—2419.

Hempel, C. 1942. "The Function of General Laws in History." *Journal of Philosophy* 39:35—48.

Hitchcock, C., and J. Woodward. 2003. "Explanatory Generalizations, Part I: a Counterfactual Account." *Noûs* 37:1—24.

Jones, B., F.Baumgartner, C.Breunig, et al. 2009. "A General Empirical Law of Public Budgets: a Comparative Analysis." *American Journal of Political Science* 53:855—873.

King, G., R. Keohane, and S. Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Levy, J.S. 1988. "Domestic Politics and War." *Journal of Interdisciplinary History* 18:653—673.

Lieberson, S., and F. Lynn. 2002. "Barking up the Wrong Branch: Alternatives to the Current Model of Sociological Science." *Annual Review of Sociology* 28:1—19.

Lipset, S. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53:69—105.

Narang, V., and R.M. Nelson. 2009. "Who are these Belligerent Democratizers? Reassessing the Impact of Democratization on War." *International Organization* 63:357—379.

Ostrom, E. 2005. *Understanding Institutional Diversity*. Princeton: Princeton University Press.

Pape, R. 2005. *Dying to Win: The Strategic Logic of Suicide Terrorism*. New York: Random House.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Przeworski, A., M. Alvarez, J. Chelbub, and F. Limongi. 2000. *Democracy and Development: Political Institutions and Well-being in the World, 1950—1990*. Cambridge: Cambridge University Press.

Ragin, C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Ragin, C., and G. Schneider. 2009. "Case-oriented Versus Variable-oriented Theory Building and Testing." Manuscript, University of Arizona.

Ramirez, F., Y. Soysal, and S. Shanahan. 1997. "The Changing Logic of Political Citizenship: Cross-national Acquisition of Women's

Suffrage Rights, 1890 to 1990." *American Sociological Review* 62: 735—745.

Ray, J. 1993. "Wars between Democracies: Rare or Nonexistent?" *International Interactions* 18:251—276.

Rodrik, D. 2006. "Industrial Development: Stylized Facts and Policies." Manuscript, Harvard University.

Rueschemeyer, D., E.H. Stephens, and J.D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.

Russett, B. 1995. "The Democratic Peace: 'and yet it Moves.'" *International Security* 19:164—175.

Russett, B., and H. Starr. 2000. "From Democratic Peace to Kantian Peace: Democracy and Conflict in the International System." In *Handbook of War Studies*, 2nd edition, edited by M. Midlarsky. Ann Arbor: University of Michigan Press.

Ryckman, M., and G. Goertz. 2009. "Rethinking and Re-estimating the Impact of Wealth(GDP/capita) on Civil War Onset." Paper presented at Peace Science Society Meetings.

Sambanis, N. 2008. "Terrorism and Civil War." In *Terrorism and Development*, edited by P. Keefer and N. Loayza. Cambridge: Cambridge University Press.

Seawright, J. 2002. "Testing for Necessary and/or Sufficient Causation: Which Cases are Relevant?" *Political Analysis* 10: 178—207.

Sikkink, K., and C.B. Walling. 2007. "The Impact of Human Rights Trails in Latin America." *Journal of Peace Research* 44: 427—445.

Singh, S., and C. Way. 2004. "The Correlates of Nuclear Proliferation: a Quantitative Test." *Journal of Conflict Resolution* 48: 859—885.

Solingen, E. 2008. "The Genesis, Design and Effects of Regional Institutions: Lessons from East Asia and the Middle East." *International Studies Quarterly* 52:261—294.

Tomasky, M. 2008. "How Historic a Victory?" *New York Review of Books* 55:44—47.

Woodward, J., and C. Hitchcock. 2003. "Explanatory Generalisations, Part 1: a Counterfactual Account." *Noûs* 37:1—24.

Young, J. 2008. "Repression, Dissent, and the Onset of Civil War." Ph.D.dissertation, Florida State University.

Zorn, C. 2005. "A Solution to Separation in Binary Response Models." *Political Analysis* 13:157—170.

16 范围

一项研究中的被试构成了某个假想的超总体的一个随机样本，这个旧观点仍值得片刻敬默。

——戴维·弗里德曼(David Freedman)

16.1 导言

由于“范围”一词通常不出现在方法类著述中，先简要介绍它是有益的。来自自然科学的一个简单例子恰好提出了模型拟合和模型设定问题，它将是本章的关注点。该例就是物理学的胡克定律(Hooke's law)，即弹簧上的应变(strain)和应力(stress)成比例(Freedman, 2009把它作为核心例子)。如果在一个弹簧上悬挂一个重物$_i$，该弹簧的长度为长度$_i$，那么该定律为：

$$长度_i = \beta * 重物_i + \varepsilon_i \tag{16.1}$$

在(如拉普拉斯和高斯的)古典物理学中，ε_i 指测量误差，是测量工具、测量者和其他因素的组合，该组合使得观测值有别于真实值。

依照现在的情形，等式(16.1)没有范围限定。因此，可以推定胡克定律对所有的重量以及在宇宙中所有的空间场域中都适用。只有当分析者在该定律的适用性上强加一个或多个限定时，才引入范围条件。

针对胡克定律，有两个自然的方法可引入范围限定。第一种是断言该定律只对一定范围的重量才有效。正如物理学家所知，很多"科学定律"在极端条件下会失效。在这种情况下，大重物的使用要求分析者在等式(16.1)中加入二次项。因此，如上所述的胡克定律的范围仅限定在某种阈值以下的重物。在该阈值以上，人们就需要通过调整模型中的变量来修正定律。①

限定胡克定律的第二个自然的方法是如下断言，即只有当重力与地球引力相等时，该定律才有效(或者主张该定律的性质将因与地心的距离而变)。这里，人们设定了作为定律起作用的必要条件的"环境"和"背景条件"。这个环境实际上是在等式(16.1)中没有设定的一个或多个隐变量。除非这些变量显现(或者假定取特殊值)，否则定律中表述的一类关系将不适用。

由于胡克定律在应用性上受到限定，因此出现了对范围条件的需要：在各类子总体(例如，对于非常重的物体来说，弹簧起作用的方式会变)以及各种环境(例如，可替代的重力)中，模型的参数是不稳定的。事实上，没有哪个物理学理论会有普遍的适用范围。也许除了量子力学的基本理论，所有的理论和概括都对环境敏感，即有范围限定。

将范围看做这样的一组变量是很有益的：

如果范围条件 S_i 成立，处理效应就是 β。

在这些范围条件以外，我们或许知道或不知道实际的关系是什么，但是我们会怀疑或者知道关系是不同的。简言之，一个范围条件就是关于因果同质性(即某个领域)的一种主张，在该领域中，可以期望因果效应是稳定的。

不足为奇，社会科学中的理论的范围比在自然科学中更受限定。虽然胡克定律很有普遍性，因为该定律适用于大多数重物和地球上的

① 只要重量足够大，所有的弹簧最终都会损坏。这个损坏点就是一类最终分界点。当达到这个点时，规律就不适用了。

大部分环境,但是社会科学中的模型在不同的子总体和环境中是异常脆弱的。在一个子总体或一个新的总体中运行同一个模型,很可能产生不同的参数估计值。

16.2 针对因果异质性问题的模型内回应

为了避免必须限定范围,学者可以通过改变因果模型来处理模型拟合和因果异质性(causal heterogeneity)问题,即我们所说的"模型内反映"(within-model responses)。我们简要地考察其中最流行的一些模型内反映。出于我们将探索的一些原因,通过改变因果模型来解决因果异质性问题,这种选项对于定量学者来说比定性学者更有吸引力。为什么高度严格的范围边界更可能在定性研究而非定量研究中被发现,这就是原因之一。

16.2.1 定量研究中的回应

对胡克定律的讨论表明,当在定量研究中遇到因果异质性问题的时候,有一种有益的思考方式,即是使用范围条件还是使用模型内的解:

$$长度=\beta * 重量+\varepsilon,重量<S_1 \quad (16.2)$$

$$长度=\beta * 重量+\gamma * 重量^2+\varepsilon,重量<S_2 \quad (16.3)$$

对于较轻的物体 S_1,我们有一个简单的等式。如果要增加范围边界以便包括更重的物品,即 S_2,就需要添加一个二次项。通过加入二次项(即重量^2),就在模型内部解决了范围问题。

范围决策涉及我们所谓的基础权衡(Fundamental Tradeoffs)的各种形式。虽然它不是我们知道的一个宇宙定律,但至少在社会科学中几乎一定可以避免的是,当人们想要增加任意一个已知的理论(而不提

出一个全新的理论）的范围时，修正了的理论会更加复杂（Przeworski and Teune, 1970）。因此，基础权衡会追问：在范围上提升的概括性是否值得在简约上有损失（即复杂性增加）。只要损失的简约性造成的代价不太高（胡克定律当然不是如此），模型内的解往往就是一个好主意。

随机化并不是对范围问题的可靠解，注意到这一点很重要。举个简单的例子，假设有一个处理（treatment）（如吸毒）对男人有正效应而对女人没有效应。因此，个体的性别就是一个范围条件，它界定着这种处理得以起作用的个案范围。如果我们没有意识到这个范围，并在所有的个体身上运行一个实验或统计模型，也许假定这种处理会对每个人都有一个平均效应，而实际上没有。

一个自然的反应就是将性别作为控制变量。正如无数哲学家和方法论学者所强调的那样，增加控制变量对于产生同质的子总体来说至关重要，反过来对于好的因果推断也必不可少。控制变量是我们所谓的模型内反映的一个经典例子，因为将范围变量纳入模型会使模型更加复杂。因此，研究者可能有如下模型：

$$Y = \beta_0 + \beta_1 T + \beta_2 S + \varepsilon \tag{16.4}$$

在这里，当控制 S（性别）时，可以检验 T（治疗）的效应。在这个模型中，我们将确切地看到，被试的性别是模型的一个重要部分。

众所周知，引入控制变量能够彻底改变所研究的自变量（即 T）的参数估计值。在范围情境下，用控制变量可检验子总体内部的因果同质性，该子总体是由控制变量的不同值界定的（一个漂亮的直观解释，参见 Berk, 2004：第 1 章）。如果加入的控制变量改变了参数估计值，研究者就会怀疑 X 在整个总体中并没有恒常的效应。在我们的简单例子中，当加入性别变量后，β_1 会改变。这意味着，处理效应因性别定义的子总体的不同而变。

加入控制变量可改变参数估计值 β_1，这表明因果异质性是一个问题。然而，加入控制变量本身并不能解决问题：它是一种诊断工具，能告诉研究者因果异质性是一个问题。

为了捕捉异质性，一个自然的举动是考虑性别和处理的交互项。

交互项的系数可真正说明处理对性别的依赖性。它让我们看到处理效应完全在男性中起作用。用这种方式加入交互项，使模型变得复杂，这是针对因果异质性(causal heterogeity)的模型内解(within-model solution)的一个好例子，它停止了直接施加范围条件。[①]在不必将分析仅限定在 S 上的一个特殊值的情况下，可以建立模型来研究 X 的效应如何依赖于 S。如果潜在的范围条件仅与一个自变量有交互作用，而且这个影响可以用一个简单的交互项来正确地建构，那么这种解就很有道理。

一种不同的方法是，在由一个范围变量所界定的子总体中单独估计整个模型。例如，研究者可以分别针对男性和女性来估计 T 的效应。如果研究者相信，因果机制在不同的子总体中也相当不同，那么这个解就是一个好的选择。例如，在关于国际关系冲突的文献中，分析者有时候针对不同的时代分别估计模型，例如，第一次世界大战之前、第二次世界大战之后和冷战后(例如，Senese and Vasquez, 2008)。就跨越不同时期的很多变量来说，它们的参数估计值的变动通常相当大。

总之，我们认为，针对因果异质性问题有三种常见的模型内回应，它们阻止对范围进行限定：(1)控制变量，(2)交互项，(3)在子总体中估计整个模型。(1)和(3)这两个解实际上是诊断工具，用它们可以确定因果异质性问题的存在，从而需要限定范围。(2)这个解原则上可以在不施加范围限制的情况下，实质地处理一个因果异质性问题。

16.2.2 定性研究中的反应

对于定性研究者来说，如果不限定其论证的范围，他们如何回应因果异质性问题呢？在这种传承下，也可能用模型内的解来阻止范围限定。然而，如在定量传承中一样，这些解常常以模型的简约为代价，从而提出了基础权衡问题。

① “某种因果关系在个体、情境、处理和结果的变异中在多大程度上成立，对它们的估计从概念上讲类似于对统计交互项的检验。”(Shadish et al., 2002:86)

在定性传承中，一个可能的模型内的解涉及在初始的集合论模型中加入一个或多个额外的因果路径。例如，考虑一个定性模型，如 $Y = AbC + BCD$。为了容纳更多的个案，研究者或许会增加一个新的因果路径。因此，初始模型可能被修改为 $Y = AbC + BCD + EF$。为新个案增加新路径在多大程度上是值得的，这取决于各种考量，包括在研理论的性质以及新路径在多大程度上适用于不止一个新个案。研究者不想对每一个新个案都必须增加一条新路径。

一个模型仅在一个给定的情境中起作用，这个发现也能促使研究者在初始模型的因果路径上加入新变量。例如，再考虑一下集合论模型 $Y = AbC + BCD$。假设研究者发现，这个模型仅适用于那些具备某个特定特征 Z 的单元。研究者可以通过使其成为一个必要条件，将这种特征当成模型的一部分：$Y = Z * (AbC + BCD)$。除非 Z 出现，否则研究者不能期待 AbC 或 BCD 对于 Y 来说是充分的。

对于那些不存在 Z 的个案来说，研究者需要确定能产生 Y 的一套不同的原因条件。这些新的原因可以与一些理论联系在一起，这些理论可以与初始模型中的理论关系不大。例如，研究者或许发现 $Y = zEF$。这些新变量 E 和 F 也许和来自初始模型的变量无关。然而，新模型可以与初始模型结合起来，以避免一个范围限定。最终的结果模型将是不太简约的：$Y = ZAbC + ZBCD + zEF$。

如此看来，无论在定量还是定性研究中，原则上讲，都可以通过修正因果模型来避免限定一个理论的范围。如果已有适当的修正并且它们有吸引力——即基础权衡是值得的，那么学者就没有什么理由去诉诸范围限定。另一方面，如果模型内的改变是不可知的，或者引入了难以处理的复杂状态，范围限定就成为一个可理解的替代了。

16.3 为什么利用范围条件？

有时候，学者能通过修正他们的因果模型来避免使用高度限定性

的范围条件，尽管如此，他们最终经常对其模型的范围或明或暗地施加了一些限定。在本节中，我们考察因果复杂性的存在以及对数据拟合性的关心如何导致学者使用了范围条件。

16.3.1 因果复杂性

作为一般规则，我们建议：

> 如果扩张一个因果模型的范围会要求该模型变得相当复杂，那么施加范围限定就变成一个有吸引力的选择。

该提议直接出于这样的观点，即在增加一个因果模型的普遍性和使该模型达到最简约之间存在着一种权衡。当分析者判定扩大模型的普遍性所带来的收益不抵使模型简约带来的损失时，就经常使用范围限定，这种简约性与必须引入的复杂性增加有联系。实际上，分析者通常不知道如何修正模型使之涵盖更广泛的个案。那么，范围限定就成为设定模型据以运行的领域的必要工具。

在实践中，范围限定在社会科学中往往非常模糊或者完全不明确（在定性和定量研究中皆如此）。学者经常含蓄或偶尔明确地在时间段或区域上使用范围限定。在操作的意义上，将范围限定在一个具体的时间段或区域上，这是相对精确的（例如范围仅限于非洲），但是在理论意义上并不精确。理想上讲，我们可以跨越具体的时空来确定抽象的范围条件。

为什么范围限定在区域和时间段上是被模糊设定的，原因之一在于，许多重要的东西会随着时间和区域的变化而变。例如，非洲在很多方面不同于拉丁美洲，如气候、人口密度、文化、殖民历史等。研究者也许不知道其中哪些具体的差异最重要，以及为什么它们恰好重要。因此，当研究者相信范围仅限于非洲乃是根本，他或她就不能确认哪些复杂的原因使得该范围是恰当的。模型内的解之所以无用，是因为它们要求已经解决了这些理论复杂性。

所设计的定性因果模型要适应总体内的每一个个案，这个事实有助于解释定性学者为什么比定量学者更容易使用范围限定。就一个集合论模型来说，加入一个或少量新个案也许会要求分析者从根本上改变初始模型。得到的因果模型的简约性可能大降。此外，学者也许甚至不知道通过哪些方式改变模型可以使之容纳新个案中的因果模式。

举一个假想的例子，一位定性学者不但需要知道当加入额外的个案时 D 是一个重要的变量，还要知道这些个案缺失的因果组合是 BCD。斯考切波(Skocpol, 1979)在其著名的社会革命理论中，将范围限定在非殖民地国家。对于她来说，扩大范围以便包含后殖民地的社会革命，如墨西哥和伊朗，这将要求加入一些新变量。此外，她将面临的任务是将这些新变量加入她的理论中，以便使所有项共同起作用。得到的理论有复杂性，这种复杂性将不会产生那个使斯考切波出名的相对高雅的结论。因此，当斯考切波的确尝试解释后殖民国家的社会革命时，她提出了一个独立的理论(Goodwin and Skocpol, 1989)。

16.3.2 更好的拟合

在定量研究中，分析的目标通常是估计所研究的个体变量的效应。就此而言，一般要施加范围条件来处理与因果异质性有关的问题。但是在定性研究中，目标是拥有能够达到与数据高度拟合的因果模型，范围条件常常与此目标有关。研究者引入范围限定的目的是提高模型的整体拟合度。

表 16.1 用一种简单的方式展示了这种对提高拟合度的关注(关于这种观点的拓展分析，参见 Ragin and Schneider, 2010)。在这个表中，我们看到两个二分变量之间的关系：高的人均 GDP(自变量)和民主(因变量)。从集合论的视角看，拟合度总体而言相当好：高的 GDP 对于民主来说几乎是充分的。然而，有 8 个个案有高的人均 GDP 却没有民主，因此违反了集合论关系。理想情况下，定性研究者希望有 0 个反例，即低权力的单元应该为空。

有一种提高拟合度的方式，即探索这些“问题”个案是否存在一些

可作为范围条件的基础的共同之处。如果是这样，就可以将它们从分析中消除，留下较强的、更清晰的结果（即一个完美的充分条件关系）。结果表明，这些问题个案（除了新加坡）都严重依赖于石油，例如，波斯湾地区的石油国家。因此，如果研究者引入的范围限定排除了所有的高度依赖石油的国家，就会显著地提高模型的集合论拟合度。

表 16.1　宽范围：人均 GDP 和民主

		高的人均 GDP 0	高的人均 GDP 1
民主	1	55	37
	0	44	8

$\chi^2 = 9.5$，$p = 0.002$，$N = 144$

1995 年

资料来源：Gerring，2007。

研究者也许倾向于相信，这种范围限定只不过是消除一些（依据某种模型所界定的）奇异值罢了。然而，范围限定是一个抽象的变量（例如高度依赖石油），而不简单是一系列具体的国家（如沙特阿拉伯、也门）。此外，在检验高度人均 GDP 对民主的影响时，可能有好的理论理由来说明为什么要去掉高度依赖石油的国家。如果研究者相信，高度人均 GDP 是诸如“经济发展”这样的宽口径概念的衡量指标，那么，加入这些个案就会增加对测量误差的关注，即石油依赖型经济也许有高度人均 GDP 却没有真正的经济发展。

另外一个方式是证明石油依赖型国家与几乎所有其他国家都不同，因为其税收不依靠向公民征税。如果有人认为，高度 GDP 通过什么机制导致民主，这与社会中的国家基础权力及展现有关，那么就有很好的理由排除这些个案，因为它们没有显示所提出的机制。

为了进一步探索这个想法，我们收集了 1995 年的石油出口数据，并去掉了所有高度依赖石油的国家。得到的全体个案产生了表 16.2。即使新加坡仍然出现在右下角单元格中，关于充分关系的集合论发现还是相当明显的。从定性研究的立场看，可以将它看成是一种受理论

驱动的范围限定，这种范围限定以一种在方法论上适当的方式来处理问题个案。在权衡方面，我们以相对次要和理论上站得住脚的范围变化为代价，得到了一个近乎完美的充分关系。

表 16.2　窄范围：人均 GDP 和民主

		人均 GDP 0	人均 GDP 1
民主	1	54	35
	0	40	1

$\chi^2 = 19.1$, $p = 0.000\,1$, $N = 130$

1995 年

当我们已经聚焦在充分条件单元格时，一种范围条件的应用会潜在地从所有单元格中移除个案，注意到这一点很重要。在本例中，范围条件从所有 4 个单元格中移除了共 14 个个案。从集合论的立场看，这不是一个大问题，因为右上角的单元仍被很好地占据着。如果范围条件移除了右上角单元的所有（或几乎所有）个案，关系将仍然是一种接近充分性的关系，但它会变得微不足道（Goertz，2006；Ragin，2008）。

从定量的角度讲，研究者也可以就表 16.1 追问范围限定的使用问题。如果范围条件运行良好的话，会发生两件事情：N 下降，拟合度上升。由于 N 下降，显著性水平会受影响，但是希望拟合度的上升能充分补偿那种损失。在本例中发生的是：我们获得一个更有实质意义的、更具有统计显著性的 χ^2。通过比较表 16.1 和表 16.2 可以看到这个结果。N 下降了 14，而 χ^2 从 9.5 上升到 16.2，并且显著性水平也提高了不少。

更大的问题仍然存在：值得这么做吗？在大样本统计分析中，答案往往是否定的。χ^2 的增加固然好。①然而，其代价却是概括范围的降低和（看起来不那么容易）获得相关数据所耗的时间。研究者将不得不辩解为什么使用范围条件而不是一个模型内的解。施加的范围限定能否

① 考虑到 χ^2 对空单元是敏感的，随此数据而增加的发生比率会变得非常大。

抵住模型拟合度方面的收益，这取决于在表述这种辩解时的难易程度。

相比之下，在定性研究中，通过范围限定可以消除少量问题个案，从中获得的收益往往超过了普遍性上的损失。当与定量研究者相比时，定性学者寻求更强的概括性（即理想条件下没有例外）。这种传承的目的是发现强的集合论关系，该目的时常表明减少一个总体的范围，使之仅包括少量个案是正当的。

16.4 范围和经验检验

有时候，学者会明确地论证，他们的理论适用于所探究现象的所有情况。例如，在国际关系冲突的文献中，我们看到一些名家认为他们的理论几乎适用于所有地方。

> 大多数理性选择理论家（包括结构实在论者）都不声称他们的理论应该受限于时空，因此期望在政治上的相关两方组（dyads）或者其他子集合中发现的关系也在所有的两方组中成立。事实上，在形式的理性选择理论家中间，布埃诺·麦斯奎塔（Bueno de Mesquita，1981）明确地指出，关于战争的期望效用理论应该适合于所有的地区和时期。肯尼思·华尔兹（Kenneth Waltz）在一部数学表述较少的著作中类似地指出，（与国内事务因素相反）系统结构的限制和诱导随时间推移对所有国家都有相同的影响。（Bennett and Stam，2000:555）

虽然一些学者在理论层次上暗含一个普遍的适用范围，但是在理论的提议范围（proposed scope）和经验范围（empirical scope）之间会有脱节。经验范围指的是由经验检验和探究所确立的一个模型的范围。在社会科学中，范围限定通常产生于经验上的点播，这样可以考量理论在哪里适用，在哪里不适用。这种做法无可厚非；事实上，它在自然科

学很常见，也是科学研究的关键部分。在增长性的研究纲领中，在经验发现与范围精炼（以及理论的其他方面）之间常有一种自然的来回摆动。研究者在建构与他们的理论相关的整个总体时需要一个过程，而经验结果和发现对于该过程来说是关键（Ragin，2000）。

遗憾的是，在理论的提议范围和理论的经验效度之间经常有逆关系。概括地说，我们认为，与范围比较受限的理论相比，范围较广的理论更容易在经验上是错误的（或者不能被强检验所确证）。因此，学者面临着范围和经验效度之间的基础权衡（Przeworski and Teune，1970）。虽然他们试图阐述具有高度普遍性的理论，但是他们经常必须牺牲普遍性以获得有效的解释。

在定量传统中，当讨论统计模型中大量使用的控制变量的时候，更大的普遍性和经验效度之间会出现张力。尽管使用控制变量是一个模型内的解，它帮助学者避免限定他们的理论范围，不过近期的文章提出了这样的问题，即加入多个控制变量时，研究者可以在多大程度上获得有效的解释。例如，阿肯建议，在一个统计模型中应该包含的自变量大约只有三个：

> 三个自变量准则（A Rule of Three，ART）：包含的解释变量超过三个的统计设定是没有意义的。ART 听起来很苛刻，但是事实上，它只是可靠的智识罢了。自变量数超过三个时，就没有人能进行审慎的数据分析以确保模型设定是准确的，同时确保诸多假定如研究者所宣称的那样拟合的好。（Achen，2002：446）

如果研究者认真对待这个建议，那么范围限定就成为一项有吸引力的选择。虽然范围限定约束了普遍性，但是它们却允许研究者运行一个包含较少控制变量的统计模型。不足为奇的是，阿肯倡导在小范围内进行更集中和高质量的分析：

> 依然更加重要的是，大的、模糊的线性回归和概率单位方程（probit equations）似乎需要很多控制变量，恰恰是因为它们将不

> 在一起的各类观测项塞在一起。国家、战争、种族类别、宗教偏好、教育水平和其他可以改变人们的系数的变量用虚拟变量来“控制”，这些虚拟变量完全不足以建构它们的效应模型。结果是一长串自变量清单、一袋子乱七八糟几乎无关的观测项，以及通常带有无意义（但在统计上是有几个星号显著的）结果的无希望的糟糕设定……反而，专业的研究习惯需要更强调在定量社会科学中产生的众多经典技巧：图、列联表以及只简单考察数据。这些方法很简单，不过它们是精致地简单。在我们正使用的精致统计工具中有诸多假定，这些技巧往往揭露了这些假定是失败的，因此挽救了我们的推理错误。做这样的工作是缓慢的，它要求我们限定在有少量解释因素的情形——通常不超过三个。将我们限定在数据子集（其中的假定是合理的）的同时也使我们限定在仅需要有少数解释因素的个案了，因此，我们可以进行创造性地思考，这是科学的精髓所在。因此，限定在同质的数据子集（及其概率单位和 logit 的对等项）的小型回归设定远远不是一个限制，而恰恰是我们进步的最佳机会所在。（Achen, 2005:337, 338）

这种论点与定性学者是相投的。在使用涵盖广泛时空领域的研究设计时，定性研究者犹豫不决，这种状态十分符合上述论点。

在统计传承中，匹配方法虽然在此背景下不被经常提及，但其运用涉及降低经验范围，以便进行更好的经验检验。当使用匹配方法的时候，定量学者会放弃一些个案，更关注其他个案。因数据性质的不同，可能要放弃相当大比例的个案，因为它们不匹配。包含在匹配中的个案并不是总体中所有个案的随机子集。其结果是，通过匹配生成的结果的经验范围比在总数据集合中执行标准统计分析的范围要小。

实验法的使用也经常涉及限定普遍性来执行强的经验检验。这种权衡可以通过内部效度与外部效度的语言来讨论。对于实际研究的被试总体来说，虽然设计合理的实验可以达到高水平的内部效度，但是在将这些实验结果向更广的情境推广时，研究者经常感到有难度。实验结果的外部效度通常是一个真问题（Morton and Williams, 2010）。虽

然这些问题未必无法克服,但它们往往向实验结果的普遍性提出了挑战。

当然,在定性传统中也一直存在同样的问题和挑战。例如,个案研究者面临一个完全标准的问题,即研究结果的普遍性何在。虽然对于一个个案的解释可能有说服力,但是学术界想要知道它的适用面是否更广。在此意义上,个案研究者和实验者必须经常处理相似的外部效度问题。

16.5 结论

范围问题引发了社会科学研究中的基础权衡。一种权衡关注普遍性和简约性之间的张力。学者在提升普遍性的尝试中,可能扩大他们理论的范围。然而,扩大范围通常需要使因果模型复杂化,从而使理论的简约性降低。牺牲简约性以获得更大的适用范围,这种做法是否合理,取决于要求的复杂性程度以及范围的扩展程度。在定性传统中,扩展范围通常是不合理的,因为即使适度的扩展也经常要对因果模型进行复杂的修正。相比之下,在定量传承中,扩展范围更经常是合理的,因为分析者有时可以通过对因果模型进行适度的调整而适应许多新的个案。

另一种权衡关注两个方面的张力,一方面是普遍性,另一方面是模型的拟合问题。范围限定如果在模型拟合或经验效度上取得了大收益,它通常就是合理的。在定性研究中,分析者少有例外,都寻求高度概括,范围限定可以显著提高模型对数据的拟合度,尤其是当限定可以在理论基础上得到明确辩护的时候更是如此。传统上讲,定量学者已经不愿意降低他们的论证范围以获得更好的数据拟合。然而,近年来,一些定量方法论者更多鼓励使用范围限定。在定量研究中,牺牲普遍性以进行更“强”的检验并提高效度,这可能是一个新趋势。

参考文献与推荐读物

有关范围问题的讨论,严谨的方法论著述出奇地少。最有名的例外可能是科恩(Cohen, 1989),沃克和科恩(Walker and Cohen, 1985)。我们在格尔茨和马奥尼(Goertz and Mahoney, 2009)中讨论了概念范围和因果范围问题。范围问题大概出现在许多重要的定性方法和研究设计著述中,如格迪斯(Geddes, 2003:95—98, 152—157);乔治和贝内特(George and Bennett, 2005: 25—27; 119—120)和拉金(Ragin, 2000:61—62)。凯泽(Kiser, 1996)强调了"抽象的"范围条件的重要性。科利尔和马奥尼(Collier and Mahoney, 1996)讨论了范围条件和选择偏误之间的关系;马奥尼和格尔茨(Mahoney and Goertz, 2004)处理了有关范围条件和负个案选择的问题。

从定量视角对范围的良好讨论是巴特尔斯(Bartels, 1996)。在定量研究中,早期的一个关于离群值和范围限定的重要陈述是博伦(Bollen, 1990)。在定量方法论中,关于通过使用简单的模型提升推断的研究,参见阿肯(Achen, 2002)以及《冲突管理与和平科学》(*Conflict Management and Peace Science*)杂志的一本专集,其中的文章包括阿肯(Achen, 2005),克拉克(Clarke, 2005),雷(Ray, 2005)和其他。本章考量的基础权衡(Fundamental Tradeoffs)则是在普沃斯基和托伊内(Przeworski and Teune, 1970:20—23)的不同背景下提出来的。

Achen, C. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5: 423—450.

Achen, C. 2005. "Let's Put Garbage-can Regressions and Garbage-can Probits Where They Belong." *Conflict Management and Peace Science* 22:327—339.

Bartels, L. 1996. "Pooling Disparate Observations." *American Journal of Political Science* 40:905—942.

Bennett, D., and A.Stam. 2000. "A Cross-validation of Bueno de Mesquita and Lalman's International Interaction Game." *British Journal of Political Science* 30:541—560.

Berk, R. 2004. *Regression Analysis: A Constructive Critique*. Newbury Park: Sage Publications.

Bollen, K. 1990. "Political Democracy: Conceptual and Measurement Traps." *Studies in Comparative International Development* 25:7—24.

Bueno de Mesquita, B. 1981. *The War Trap*. New Haven: Yale University Press.

Clarke, K. 2005. "The Phantom Menace: Omitted Variable Bias in Econometric Research." *Conflict Management and Peace Science* 22:341—352.

Cohen, B. 1989. *Developing Sociological Knowledge: Theory and Method*, 2nd edition. Chicago: Nelson-Hall.

Collier, D., and J. Mahoney. 1996. "Insights and Pitfalls: Selection Bias in Qualitative Research." *World Politics* 49:56—91.

Freedman, D. 2009. *Statistical Models: Theory and Practice*, revised edition. Cambridge: Cambridge University Press.

Freedman, D. 2010. *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*. Cambridge: Cambridge University Press.

Geddes, B. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor: University of Michigan Press.

George, A., and A.Bennett. 2005. *Case Studies and Theory Development*. Cambridge: MIT Press.

Gerring, J. 2007. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Goertz, G. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41:88—109.

Goertz, G., and J. Mahoney. 2009. "Scope in Case Study Research." In *Handbook of Case-oriented Research*, edited by D.Bryne and C.Ragin. Newbury Park: Sage Publications.

Goodwin, J., and T. Skocpol. 1989. "Explaining Revolutions in the Contemporary Third World." *Politics and Society* 17:489—509.

Kiser, E. 1996. "The Revival of Narrative in Sociology: What Rational Choice Theory can Contribute." *Politics and Society* 24: 249—271.

Mahoney, J., and G. Goertz. 2004. "The Possibility Principle: Choosing Negative Cases in Comparative Research." *American Political Science Review* 98:653—669.

Morton, R., and K.Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.

Przeworski, A., and H. Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.

Ragin, C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Ragin, C., and G. Schneider. 2010. "Comparative Political Analysis: Six Case-oriented Strategies." Manuscript, University of Arizona.

Ray, J. 2003. "Explaining Interstate Conflict and War: What should be Controlled for?" *Conflict Management and Peace Science* 20:1—30.

Ray, J. 2005. "Constructing Multivariate Analyses(of Dangerous Dyads)."*Conflict Management and Peace Science* 22:277—292.

Sambanis, N. 2004. "What is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48:814—858.

Senese, P., and J. Vasquez. 2008. *The Steps to War*. Princeton: Princeton University Press.

Shadish, W., T. Cook, and D. Campbell. 2002. *Experimental and Quasi-experimental Designs for General Causal Inference*. Boston: Houghton Mifflin.

Skocpol, T. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.

Walker, H., and B. Cohen. 1985. "Scope Statements: Imperatives for Evaluating Theories." *American Sociological Review* 50:288—301.

17 结论

差异性恰恰和相似性一样，可以成为相互吸引的原因……一定类型的差异使我们倾向于接近他人。这意味着相互补充，而非反对和排斥。

——埃米尔·涂尔干(Émile Durkheim)

金、基欧汉和维巴在《社会科学中的研究设计》一书中给出如下总结性文字："定性研究者所面临的方法论问题也是所有科学研究者需要面对的。只有内在逻辑得到充分的理解和遵循后，所有的社会科学研究才能做出有效的推论。"(King, Keohane and Verba, 1994:230)。相比之下，在本书行将杀青之际，我们希望再次提请注意定性研究和定量研究在性质上的重要差别——这些差别涉及研究设计、数据分析、概念和因果推断。在陈词滥调的相似性(例如，研究目的是通过使用系统的程序进行有效的推断)之外，没有一套原则可以将所有的社会科学研究统一起来。

然而，我们确信，在定量和定性研究范式之间存在着对话的空间。虽然两种研究范式的差异很大，但是在一项旨在解释社会和政治世界的总项目中，二者恰好可以互补。就像涂尔干关于有机团结的洞见一样，我们在定量和定性研究之间看到了富有成效的合作的现实可能性，这种可能性建立在相互尊重与欣赏之上。然而，为了实现这种可能性，需要我们正面理解和承认在两种研究传统之间存在的许多重要差异。

表 17.1　汇总的对比 I:个案

描述项	定量传统	定性传统	所在章
(1) 解释个案中的结果	很　少	常　见	1，3，4，6
(2) 跨个案对个案内	个案间为要	个案内为主	1，4，8
(3) 因果机制	可　选	必须被定义	7，8，14
(4) 过程追踪法	可　选	标准的	7，8，14
(5) 反事实分析	主要是个案间	主要是个案内	9

表 17.2　汇总的对比 II:因果性和因果模型

描述项	定量传统	定性传统	所在章
(1) 关注个体变量	标准	有时	2，6，8，9
(2) 因果构型，交互项	有时	常见	2，4，6，8
(3) 因果效应	平均处理效应	必要的和/或充分的	2，3，4，6，8
(4) 反事实的目的	阐释模型和参数估计	个案内的因果推断	9
(5) 殊途同归	内隐的，模型是路径	明确的	2，4
(6) 因果模型中的聚合	加法、对数线性、连接函数中可加	极大值、极小值、充分非必要条件的必要非充分部分(INUS)	2，4

17.1　差异概要

表 17.1 至表 17.5 汇总了本书考察的两种传统之间的关键差异。[①]这些表无意呈现所有的差异，只呈现两种传统在广度和深度上的某些

① 再次申明，我们的观点仅适用于集中关注因果推断的研究，因此我们的汇总不包括大多数诠释性分析。要想比较诠释学传统和我们描述的这种研究，就需要一个完全不同的清单。

对比。每个表都有五项，共得到包括 25 种差异的清单。在一些情况下，这两种实践的差异十分巨大，而在其他情况下则小得多，但是我们觉得在每一个情况下差异都是明显的。

表 17.3 汇总的对比 III：总体和数据

描述项	定量传统	定性传统	所在章
(1) 范围	宽	窄	6
(2) 选择个案研究	代表性、随机	(1，1)单元格最重要	14
(3) 选择因变量	不　选	有时选	14
(4) 数据格式(如数据表)	行是个体观测值	行是变量的构型	2
(5) 三角数据	异方差性	必要或充分条件	2

表 17.4 汇总的对比 IV：概念与测量

描述项	定量传统	定性传统	所在章
(1) 术语	变量-指标	概念-数据	10
(2) 本体论	未观测的变量引起指标	定义概念的诸维度	10
(3) 变异	所有的变异都有优先重要性	区域中的数据变异不改变意义	11
(4) 变量转换的逻辑	偏态性、统计模型中更拟合	语义学和意义转换	11，12
(5) 类型学	互斥	重叠或互斥	13

表 17.5 汇总的对比 V：不对称性

描述项	定量传统	定性传统	所在章
(1) 解释 0 不同于解释 1	不	有时	5，13
(2) 概念及其对立面	针对概念及其对立面使用同一个概念	经常使用不同的概念和测量	5，13
(3) 反事实 $x_i \rightarrow x_j$ 不同于反事实 $x_j \rightarrow x_i$	内含相同	通常不同	9
(4) 2×2 表，互换(0，1)和(1，0)单元	几乎所有的相关测度都等值	不同，因为一个是必要条件单元，另一个是充分条件单元	2，5，15

使用这些表的一种方式是将其作为对研究(包括研究者自己的研究)进行分类的一套“身份”检验表。如果我们关于两种传承的论证可以很好地描述研究实践,那么单个研究计划应该趋于归入两种传承的一种或另一种所提出的项目中,混合的反应相对较少。在混合方法研究中,也可能区分出哪些是定量部分,哪些是定性部分。

17.1.1 个案

我们的核心论点是断言定量和定性研究者以相当不同的方式来处理个案。在定量研究中,分析者很少解释为什么特定的个案有特殊的结果。相反,这些研究者聚焦于跨个案分析以及作为一个整体的更大总体的特征。在一项优秀的定量研究中,学者无须开展任何过程追踪或者识别因果机制就可以达致主要的研究目标。如果实施反事实分析,它也源自跨个案比较。

在定性研究中,分析者有一种相当普遍的做法,即努力解释为什么特殊的结果发生在特定的个案中。在这种探究模式中,学者在推断时大大依赖于个案内分析,时常也中等程度依赖于跨个案分析。于是,他们几乎总是使用过程追踪法,并且在所研究的特定个案中锁定背后的机制。他们也执行反事实分析,从中可以再现一个或多个特殊个案的历史。

17.1.2 因果性和因果模型

关于因果性和因果模型,定量研究者通常典型性地聚焦于个体变量的影响,以及仅在需要时加入交互项。他们依据平均处理效应来理解并定义因果关系。有时候,他们出于阐释其统计模型的目的而使用反事实,但是反事实并不被用做一种假设检验的方法。在典型的定量模型中,聚合的形式包含加法,或者至少在连接函数中是可加的。定量研究者通常不谈论殊途同归性,而是将他们的因果模型当成是代表了通向因变量的路径。

定性研究者通常更关注因果构型而不是个体变量的效应(除了必要条件以外)。他们根据必要条件、充分条件和“充分非必要条件的必要非充分部分”(INUS conditions)来理解并定义因果关系。他们经常使用个案反事实作为假设检验的一个途径。他们的因果模型普遍有这样的假定,即个案可以遵循不同的途径达到同一结果,但是这样的途径并不多。定性模型通过潜在地使用布尔运算(如取最大值或最小值),倾向于将诸多原因项组合起来。

17.1.3 总体和数据

定量研究者倾向于研究大总体,并提出涵盖广泛范围的概括。当他们选择个案研究时,会尽量选择能代表更大总体的个案。在这种传统中通常有一个不好的主意,即基于因变量的值来选择个案。定量学者将数据汇集到标准的矩形数据表中,行表示单个观测值,列表示变量。当定量分析者面对一个三角数据集时,因异方差性而修正它就理所当然了。

相比之下,定性研究者倾向于研究少量个案,并在小范围内概括。他们对感兴趣的结果和原因都存在于其中的个案特别关注。有时,他们选择个案恰恰是因为它们在因变量上的取值。在定性传统中,将数据集合的行理解为是变量值的逻辑配置(logical configurations)。当接受过定性比较分析(QCA)训练的学者在检查三角数据时,他们自然会把它解释为代表了必要条件或充分条件。

17.1.4 概念和测量

在定量传统中,测量问题一般用变量和指标术语来讨论。这里通常有一个假定,即潜变量引起它的指标,因此后者与前者相关联。在这种传统中,人们对研究和解释变量的整个变异域很感兴趣。通常因充分的统计理由(如修正偏态)而执行变量转换。当利用类型学的时候,它们指的是互斥的类别。

在定性传统中，解决测量问题通常需要探索概念和数据之间的关系。测量是一个语义学问题，它需要设定概念的界定性维度。定性学者经常假定，一个变量的特定领域的变异（尤其是变量的下限和上限值）可能与一个概念的测量不相关。定性学者很警惕变量转换，除非它们维持或增加了所探究的概念的含义。当运用类型学时，它们可能是互斥的，或者允许在多个类别中有重叠的成员资格。

17.1.5 不对称性

定量学者提出了对称的因果论证，其中用同样的变量和模型来解释一个结果是否出现。这些学者也对称地看待一个概念及其对立面，从而使得否概念（例如，不发展）等同于对立概念（例如，不发达）。他们同样潜在地假定有一个对称的反事实观，其中一个方向的变化（例如，从威权主义到民主）貌似是另一种方向的变化。最后，大多数 2×2 表的统计测度都是对称的。

相比之下，定性学者经常提出不对称的因果论证，其中用不同的变量和模型来解释一个结果是否出现。他们经常假定一个概念及其对立面是不对称的；它们可能需要有不同的定义和测度。他们同样坚持不对称的反事实观，其中一个方向的变化不见得像另一个方向的变化。最后，在定性传统中，考察 2×2 表是因为其拥有不对称性，尤其是其必要条件和充分条件有典型的模式。

17.2 社会科学中的方法论多元主义

定量和定性范式之间存在的差异未必是社会科学中矛盾的来源。在我们列举的差异中，没有任何一种意味着矛盾。事实上，一旦人们考虑到研究范式之间对比性的目标和目的，这些差异就相当易懂。从各自的目标和目的出发，两种传承都是“可以理解的”。

既然两种传承都是相对一致的系统,很多研究者强烈地倾向于一种(而不是另一种)传承也就不足为奇了。幸运的是,合作和互相尊重不要求所有学者同时成为两种传承的正式成员。没有理由反对劳动分工的存在,其中一些学者运用装备精良的方法和工具从事专业任务。

然而,两种传承都是有渗透性的、边界松散的系统,它们相互影响,这也是实情。定量和定性分析不能也不需要在相互隔离中追求各自的研究。不如说,一种传统中的方法和发现可以通过许多途径以有益的方式进入另一种传统中。而且,我们确信,混合方法研究常常是一个可行的选项。

研究者在多大程度上可以混合两种传承,其方式多种多样。在一些个案中,研究者也许主要是定量的,或者定性的,但是从另一种传统中借鉴了优秀的观点和工具。就像我们强调过的那样,通过理解另一种传承中的学者怎样做研究,我们可以学到的东西太多了。并且对于一位特定的学者来说,在没有完全接受另一种传承的情况下,当然可能在自己的研究中引入一定的实践和程序。

对于比较彻底的混合方法研究学者来说,他们甚至也促成同时并且全面地追求定量和定性研究两方面特征的目标。在这里,研究者将估计特定的变量在大总体中的效应,在该总体内解释特殊个案的特定结果。为了达到前一个目标,他或她会使用跨个案分析、统计模型和定量范式的工具。为了达到后一个目标,他或她会追求个案内分析,提出构型式的因果模型(configurational causal models),利用定性范式中的全部资源。因此,最终成果会包含两种相当不同但无矛盾的结果集合。

如果赞同一定的劳动分工和将两种传承混合的可能性,我们就拥有了社会科学的多元视野。如此看来,应该将一处重要且受尊重的空间留给定量研究、定性研究和各种混合方法研究。我们相信,如果说在这种方法论多元主义勃兴之途上仍有障碍的话,那么主要的障碍仅仅是不能清晰地辨别定量和定性分析在目标和程序上的差异,尽管二者同样正当。本书将这两种范式看做备选的传承,在阐明这些差异的同时,也培养两者之间的建设性对话。

附　录

政治科学和社会学中的一些重要论文都践行了诸多方法论实践，本附录呈现了针对这些实践的调查结果。这项调查追踪了学者在生产经验研究时实际做了什么，这些研究被相关学科认定是优秀的作品，在重要期刊上发表。至于在这一工作中典型地使用的实践是否代表“最佳实践”，虽然我们可以就此进行争辩，但是它们就是这样一些程序，学者用这些程序在各个学科内生产出最佳的实质性成果。

对发表在政治科学和社会学顶级期刊上的论文进行分层随机抽样，得到的结果如下面各个表所示。样本框是 2001 年至 2010 年发表在六本权威期刊——《美国社会学杂志》、《美国政治科学评论》、《美国社会学评论》、《比较政治学》、《国际组织》、《世界政治》——上的文章。样本依据期刊和时间段（2001—2005 年和 2006—2010 年）进行了分层。每一层选出 40 篇论文，其中有 18 篇被编码，导出总共 216 篇论文。样本中去掉了评论性文章和非经验性的理论文章。西北大学的两名博士候选人凯伦尼莎·莫罕达莉和克里斯托夫·纽伦完成了编码工作。带有全部数据的电子表可供读者索取。

表 A.1 提供了我们的文章样本的基本统计量。可以看出，定量方法论构成了绝对多数（72%对 31%）。这里重点要记住的是，有两本期刊——即《美国政治科学评论》和《美国社会学评论》——很少发表定性研究成果。其他四本期刊发表了相当比例的定性研究文章。

数据表明，这些期刊几乎不发表明显的多元方法研究成果。如何满足期刊文章长度的限制，这是开展多元方法研究遇到的一个挑战。然而，在一些子领域和期刊，比如《国际组织》（*International Organiza-*

tion)中，我们观测到这样一种传统，即在发表的定量论文中包含几个短小的个案研究。而且，近期关于比较政治学和国际关系学的著作都明确地转向了多元方法研究。

表 A.1　使用的方法论

方法论	百分比(%)
定量研究	72
定性研究	31
明确的多元方法	1
诠释性研究	2
理论的/博弈论	8

注：论文可使用不止一种方法论。

表 A.2　使用的定性方法论

方法论	百分比(%)
个案研究($N=1$)	27
小样本研究($1<N<10$)	63
中样本研究($N>9$)	8
类型学	26
定性比较研究(QCA)	1
殊途同归性	13
反事实	15
新概念	31
明确的过程追踪	22
自变量的中位数	2

注：论文可使用不止一种方法论。

在我们调查的论文中使用了各种定性方法论，表 A.2 提供了关于这些方法论的基本信息。超过 90%的定性论文都是个案研究或小样本研究，这并不奇怪。其中，小样本研究比个案研究更多一些。数据还显示，中样本研究不常见；只有 8%的定性论文有 10 个或更多个个案。

定性研究的一个很常见的部分是使用类型学，在所有论文中，大约有 1/4 包含有明确的类型学。由于定性学者不受限于需要大量个案的

数据，因此，无论出于描述性目的还是解释性目的，他们都能更容易提出类型学。这些类型学通常要引入新概念，这在定性研究中不是不常见（在全部定性论文中，有31%引入了一个新概念）。

明确地讨论反事实和殊途同归性的论文比例比较低（分别是15%和13%）。正如我们已经讨论过的，在定性研究中，反事实的使用常常是隐性的，而且不直接作为一种推断方法来讨论。反过来，这与大多数定性研究者在使用方法论工具上的非系统性有关，同样，在一项分析中殊途同归性也常常是隐含的，尽管它的有限使用更有可能与大多数定性研究的小样本有关。个案研究几乎从定义上就决定了它只能考虑一种路径。当有一个或多个个案时，殊途同归性就真的开始起作用了。

尽管在定性研究中常常暗中使用过程追踪法，该方法通常也不直接使用。在我们的样本中，只有22%的文章明确地使用了过程追踪法。

定性方法论领域的一个重大挑战是鼓励学者更多地意识到方法论问题，以及更明确地了解他们用以推断的程序。我们希望本书帮助定性研究者提升更大的方法论自觉。

表 A.3　使用的定量方法论

方法论	百分比（%）
最小二乘法（OLS）	23
Logit/概率单位回归（Logit/probit）	37
时间序列	2
面板/时间序列截面（Panel/TSCS）	18
交互项	18
讨论 R^2	6
贝叶斯方法	3
实验法	5
工具变量	3

注：论文可使用不止一种方法。

现在转向定量研究。表A.3表明，在政治科学和社会学中，主流、经典和统计的子传承仍占主导。如果我们统合常规最小二乘法（OLS）、

logit 回归(logit regression)、时间序列分析和面板方法论,就会囊括所有定量论文的 80%。[①]尽管贝叶斯技术很少出现在我们的调查中,但是该技术的使用正在上升,它可能是我们考察过的期刊中的一篇作品。同样,尽管潜在结果架构(potential outcomes framework)——又名 Neyman-Rubin-Holland 模型——在方法论文章中非常有影响力,但是在我们调查的 216 篇文章中,没有一篇可觅其踪迹。另一方面,可以将工具变量和实验法的使用看成是潜在结果架构的指标,它们合计占到所有文章的 8%。

各种传承总是混含着久存的实践和速变的实践。定性传承、定量传承和它们的子传承都无例外。我们的调查反映了过去十年的研究,但是,如果查看一个长时段,很多编码的项目得分就会发生重要变化,而且更重要的是,一些项目会加入,另一些项目会被清除。除了随时间而变,传承也随空间——可能是子领域、学科或者地理范围——而变。在比较研究和国际关系研究中,多元方法分析看起来越来越常见并享有声望,但是在关于美国政治的研究工作中远非如此。定性比较分析(QCA)在社会学中比在政治科学中得到了更多的关注和尊重。如果在欧洲和美国之间比较其身份地位也是如此。

最后,本书在特定的时空提供了方法论实践的一个快照。两种传承之争特别适合于描绘 21 世纪初的美国情形。两种传承之间的差异十分显著,以至于很可能持续到未来。不过,我们也认为,一个成长中的学者群体将跨越边界开展研究,创造性地将来自两种传承的观点组合起来。这种交叉的研究关注的是如何超越分歧,这可能很好地代表了未来数年最令人兴奋的社会科学。

① 一篇文章可能使用多种方法论,但这些种类很少重叠。

术语对照表

aggregation techniques　聚合技术
asymmetry　非对称性
average treatment effect　平均处理效应
Bayesian inference　贝叶斯推断
Boolean methods　布尔代数法
cases　个案
causal complexity　因果复杂性
causal inference　因果推断
causal mechanism　因果机制
causal models　因果模型
causal-process observation（CPO）因果过程观测
causes-of-effects approach　先果后因研究
concepts　概念
conjunctural causation　结合因果性
constant conjunction definition of causation　因果性的恒随定义
counterfactual analysis　反事实分析
covering laws　覆盖律
critical theory　批判理论
culture　传承
data-set observations（DSOs）　数据集观测项
effects-of-causes approach　先因后果研究
error　误差
experiments　实验
Fundamental Principle of Variable Transformation　变量转换的基本原则
Fundamental Problem of Causal Inference　因果推断基础问题
Fundamental Tradeoffs　基本权衡
fuzzy-set analysis　模糊集分析
game theory　博弈论
generalization　概括
heteroskedasticity　异方差性
individual case analysis　个体个案分析
Institute for Qualitative and Multi-Method Research（IQMR）　定性与多元方法研究所
interaction terms　交互项
interpretive methods　诠释法
Interuniversity Consortium for Political and Social Research（ICPSR）　政治与社会研究校际联盟
INUS condition　充分非必要条件的必要非充分部分
Logic　逻辑学
logical AND　逻辑“与”
logical OR　逻辑“或”

measurement 测量
mechanism 机制
method of agreement 求同法
method of difference 求异法
mixed-method research 混合法研究
multimethod research 多元方法研究
multiple causation 多元因果
necessary condition 必要条件
Neyman-Rubin-Holland model Neyman-Rubin-Holland 模型
overdetermination 过度决定
Paradox of the Perfect Predictor 完美预测项悖论
potential outcomes framework 潜在结果架构
Principle of Conceptual Opposites 概念对立面原理
Principle of Conceptual Overlap 概念重叠原理
Principle of Unimportant Variation 次要变异原理
process tracing 过程追踪
qualitative research 定性研究
quantitative research 定量研究
Rorschach Principle 罗夏原则
scope conditions 范围条件
set theory 集合论
statistical methods 统计模型
sufficient condition 充分条件
translation problems 转译问题
validity 效度
variable transformations 变量转换
within-case analysis 个案内分析

译后记

我们相信，读者在诸多学术场合会发现定量研究和定性研究的学者在相互贬低。某些著名的定量学者可能声称，定性研究不是科学，某些著名的定性研究者或许认定，统计均值研究不能把握丰富的社会世界。定量研究和定性研究到底是什么关系？加里·金等学者在《社会科学中的研究设计》（陈硕译，格致出版社 2014 年版）中认为，这两种研究框架的逻辑是一样的。本书对此持反驳态度，并从研究设计、因果推断、数据分析、概念形成、个案选择、概括范围等方面详细地论证了二者之间的差异。比如，在作者看来，定量研究主要用统计学与概率论语言，定性研究却以集合论与逻辑学为基础。集合论模型处理的是不对称关系，假定不同的 X 值对 Y 产生不同的效应，而统计模型通常有这样的假定，即不管 X 的起点值怎样，X 的单位变化都对 Y 产生同样的影响。定性学者用个案内分析，就所研究的特定个案说明哪些因素会产生结果，定量学者则依据个案间分析向总体做推断。定性研究和定量研究使用的标准因果模型［方程(4.1)和方程(4.2)］差异很大，难以比较。在定性的个案内因果分析中，反事实分析居于核心，而在定量研究中，反事实分析的目的在于解释统计估计的结果，做出有关因果效应的普遍论断。在定性传统中，研究者要确定一个概念的属性，定义此概念的意义，对拥有极值的个案最有自信，定量研究则要通过区分出指标来建构概念，这些指标是由所研究的概念引起的，定量学者对有极值的个案最没有把握。定量研究者试图估计大总体内的平均效应，欢迎随机选择的个案，不会赋予特定的个案以特殊的权重；定性研究者则旨在解释特殊个案，通常对某些单元的抽样有偏好，经常关注实质重要的个

案。定量学者采用标准化和取对数值等进行数据转换，定性研究者几乎不做这样的转换，而是“校准”数据，使之符合分析者所界定的概念的意义。定性研究用过程追踪进行个案内分析，定量学者关注回归分析在产生强因果推断方面的能力，可用实验法来估计某种处理的平均效应，不过实验研究未必把握住机制。在定性研究中，强概括指的是集合论的关系，其形式大致为：“所有/没有 *A* 是 *B*”，而在定量研究中，强概括的意义是两个变量之间有强的统计相关，或者存在显著的平均处理效应。

当然，这两种研究路数也在互补。统计研究者越来越用定性个案研究来补充他们的研究，而定性研究者就少量个案做出发现的时候，也关注其是否具有普遍意义，小样本的发现可能会刺激一个范围更广的利用大样本的跨个案检验。本书对这两种研究路数之差异的论证非常细致，读者必然从阅读中获益匪浅。不过需要提醒读者注意，本书并不讨论诠释性的研究，因为它不关注因果分析，而重在阐释行为的意义或批判权力的运用。

本书原文书名为 *A Tale of Two Cultures*，无论怎样直译都未达其精髓，张磊博士建议译成两种传承，同时就个别语句的译法提供了高见。笔者曾将本书作为学生课外阅读的专业文献，邀请哈尔滨工程大学社会学系 2011 级本科生常怡、吴娜等，2014 级研究生张国锋、周莹等同学参与翻译本书的初稿，西安交通大学 2012 级朱翠翠同学帮我检查中文译稿。北京大学社会学系郭志刚教授、张静教授，南京大学社会学院风笑天教授，西安交通大学社会学系边燕杰教授拨冗惠赐荐语，责任编辑王亚丽女士为本书付出了艰辛的劳动。在此一并致以衷心的感谢！读者如能通过阅读本书，厘清定性与定量研究之差异，破除二者相互贬低的境况，夯实社会研究的方法论基础，译者便会在异化世界中暂时寻得诠释性的生命之意义，这种意义是利用本书的逻辑所推断不出来的。本书的翻译得到国家社会科学基金项目（批准号：13BSH054）和国家社会科学基金重大招标项目（批准号：13&ZD177）的支持。译文不当之处，请读者批评指正。

2016 年 4 月 6 日于西交大三村寓所

A Tale of Two Culture: Qualitative and Quantitative Research in the Social Sciences
Copyright © 2012 by Princeton University Press.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

This authorized Chinese translation edition copyright © 2016 by TRUTH & WISDOM PRESS.
All rights reserved.

上海市版权局著作权合同登记号:图字 09-2014-630

格致方法·社会科学研究方法译丛

社会科学中的研究设计

[美]加里·金　罗伯特·基欧汉　悉尼·维巴　著　陈　硕　译

混合方法研究导论

[美]约翰·W.克雷斯维尔　著　李敏谊　译

两种传承:社会科学中的定性与定量研究

[美]加里·格尔茨　詹姆斯·马奥尼　著　刘　军　译

定性研究的30项基本技能

[美]约翰·W.克雷斯维尔　著　王锡苓　译

定性研究一直被描绘成定量研究的继子，且没有礼貌，因其表面的简单分析技术和在理解和解释“个案”时的错位的热情而受阻。《两种传承》终结了这种谬论，强调定性研究的特质并逐条与定量研究对比。本书通俗易懂，向传统观点发出了全面的挑战。

——查尔斯·拉金，亚利桑那大学

格尔茨和马奥尼给出了社会科学的定性和定量范式的总体目标和假定，为读者论证为什么这两种对比性的研究路数都是恰当的，促使双方的读者都能看到另类视角，并重新考虑他们自己的研究和目标。《两种传承》明确地向关于如何做研究的基本原理提出了挑战。

——珍妮特·博克斯－斯特芬斯迈耶，俄亥俄州立大学

本书从研究设计、分析目标、因果效应、概念测量、个案选择等方面，详细论证了定量研究和定性研究的差异。原著深入浅出、言简意赅、案例丰富，是一部广受欢迎的社会科学方法论著。费孝通先生曾指出，“翻译是有拘束的创作”。依我看，刘军教授的翻译遵循了费老的教导，为学界献出了一份“信达雅”的社会科学读物。

——边燕杰，明尼苏达大学、西安交通大学

本书深入地反思了定量研究，系统地捍卫了定性研究的合理性。无论读者专攻定量研究还是定性研究，都必然会在认真研读本书中获益匪浅。

——郭志刚，北京大学

上架建议：社科研究方法

ISBN 978-7-5432-2621-0

9 787543 226210 >

定价：45.00元

易文网：www.ewen.co
格致网：www.hibooks.cn

您可以通过如下方式联系到我们：
邮箱：hibooks@hibooks.cn

微信 微博